

## ABSTRACT

Title of Dissertation: Dynamics and Chemical Shielding in a Small  
Protein from NMR Spin Relaxation Experiments

Jennifer Blake Hall, Doctor of Philosophy, 2006

Directed By: Prof. David Fushman  
Department of Chemistry and Biochemistry

NMR spin relaxation spectroscopy is a proven method for characterization of the time-scales and amplitudes of intramolecular motions in proteins. In a typical NMR relaxation experiment, the significant contributions to relaxation of the  $^{15}\text{N}$  spin are from the  $^1\text{H}$ - $^{15}\text{N}$  dipolar interaction with the covalently attached amide proton and from the anisotropic component of the  $^{15}\text{N}$  chemical shielding tensor, the  $^{15}\text{N}$  chemical shielding anisotropy (CSA). Herein I suggest novel schemes for the measurement of CSA/dipolar cross-correlation rates in proteins and for direct nitrogen-detected relaxation measurements. I also conduct a series of established NMR  $^{15}\text{N}$  spin relaxation experiments on a small protein, the B3 domain of *Streptococcal* protein G, in order to measure the overall rotational diffusion tensor of this domain, to quantify any conformational exchange type motion of backbone amides occurring in this protein on the  $\mu\text{s}$ -ms timescale, and to determine the  $^{15}\text{N}$  backbone amide CSAs in this protein in solution and quantify the effect of residue-to-residue variations in the  $^{15}\text{N}$  CSA on model-free motional parameters which describe motions on the ps-ns timescale.

DYNAMICS AND CHEMICAL SHIELDING IN A SMALL PROTEIN FROM  
NMR SPIN RELAXATION EXPERIMENTS

By

Jennifer Blake Hall

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2006

Advisory Committee:  
Professor John Tossell, Chair  
Professor David Fushman  
Professor Victor Munoz  
Professor Nico Tjandra  
Professor Sergei Sukharev

© Copyright by  
Jennifer Blake Hall  
2006

# Table of Contents

|  |     |
|--|-----|
| Table of Contents .....  | ii  |
| List of Tables .....   | v   |
| List of Figures .....  | vii |
| Chapter 1: Introduction and Specific Aims .....  | 1   |
| 1. 1 Introduction.....   | 1   |
| 1.2 Motion in Proteins.....  | 3   |
| 1.3 Chemical Shielding Anisotropies in Proteins and Peptides .....                               | 10  |
| 1.3.1 Chemical Shieldings and their Relationship to Protein Structure.....                       | 10  |
| 1.3.2 The Chemical Shielding Tensor.....   | 11  |
| 1.3.3 Chemical Shielding Tensors in Proteins and Peptides.....                                   | 11  |
| 1.3 Scope of Present Work.....   | 16  |
| 1.4 Specific Aims.....   | 17  |
| Chapter 2: NMR Relaxation Theory .....   | 22  |
| 2.1 The Liouville-von Neumann Equation .....   | 23  |
| 2.2 Hamiltonian for Spin Relaxation .....  | 24  |
| 2.3 Separation of Spin Operators and Spatial Functions .....                                     | 28  |
| 2.4 Correlation and Spectral Density Functions .....   | 30  |
| 2.5 Relaxation Rates .....   | 36  |
| Chapter 3: Methods.....  | 41  |
| 3.1 NMR Pulse Sequences for Measurement of $^{15}\text{N}$ Auto- and Cross-Relaxation Rates..... | 43  |
| 3.1.1 An Introduction to NMR Pulse Sequences .....   | 43  |
| 3.1.2 Longitudinal Relaxation, $R_1$ .....   | 49  |
| 3.1.3 Transverse Relaxation, $R_2$ .....   | 51  |
| 3.1.4 Heteronuclear $^{15}\text{N}$ - $\{^1\text{H}\}$ NOE.....                                  | 53  |
| 3.2 Spectral Processing and Data Analysis.....   | 54  |
| 3.3 Protein Expression, Purification, and Characterization .....                                 | 59  |
| Chapter 4: CSA/dipolar Cross-Correlated Relaxation Rates .....                                   | 61  |
| 4.1 CSA/Dipolar Cross-Correlated Relaxation.....   | 61  |
| 4.2 Measurement of CSA/Dipolar CCRs from Coupled HSQC Spectra.....                               | 64  |
| 4.3 Spin State Selection for Measurement of CSA/dipolar CCRs.....                                | 77  |
| 4.4 Longitudinal CSA/dipolar CCR.....  | 82  |
| 4.5 Conclusions.....   | 85  |
| Chapter 5: Overall Rotational Diffusion Tensor of GB3.....                                       | 87  |

|   |     |
|---|-----|
| 5.1 Motivation.....   | 87  |
| 5.2 Method for Derivation of the Rotational Diffusion Tensor of A Molecule from NMR Relaxation Data.....  | 89  |
| 5.3 The Rotational Diffusion Tensor of GB3 .....  | 91  |
| 5.4 Comparison with Predictions from Theoretical Hydrodynamic Models.....   | 104 |
| 5.4.1 Simple Predictions based on the Stokes-Einstein-Debye Equation of Rotational Diffusion and Empirical Relations for Proteins. ....   | 104 |
| 5.4.2 Modern Hydrodynamic Models.....   | 106 |
| 5.5 Conclusions and Discussion .....  | 109 |
| Chapter 6: Local Motion in GB3. ....  | 112 |
| 6.1 Introduction.....   | 112 |
| 6.2 Combinations of Auto- and Cross-Correlation Relaxation Rates in GB3 Can Identify Conformational Exchange.....   | 114 |
| 6.3 Picture of Local Motion in GB3 is Markedly Dependent on the Model of Overall Rotational Diffusion.....  | 117 |
| 6.4 Identification of Conformational Exchange Motion through the Field Dependence of $^{15}\text{N}$ Relaxation Rates.....  | 121 |
| 6.5 Structural Dependence of Microdynamic Parameters and Conformational Exchange Motions .....  | 123 |
| 6.6 Relaxation-Compensated CPMG Measurements Detect Conformational Exchange on Slower Timescales .....  | 125 |
| 6.7 Conclusions.....  | 127 |
| Chapter 7: $^{15}\text{N}$ Chemical Shielding Tensors in GB3 .....  | 129 |
| 7.1 CSAs from NMR Relaxation Data; a Discussion of Current Techniques and Measurements .....  | 129 |
| 7.1.1 Motivation.....   | 129 |
| 7.1.2 Discussion of Multifield Analyses of $^{15}\text{N}$ Relaxation Rates.....  | 131 |
| 7.1.3 Agreement of Diffusion Tensor of GB3 Derived from Auto and Cross-Correlation Relaxation Rates at Five Spectrometer Fields as Evidence that Relaxation Experiments Have Similar Conditions. .... | 133 |
| 7.1.4 $R'_{2\text{free}}$ versus $R_2'$ as Evidence there is no Significant Conformational Exchange Contribution to $R_2$ in Measurements at Five Fields. ....  | 137 |
| 7.2 Model independent methods for measurement of $^{15}\text{N}$ CSA from Relaxation Rates at Several Field Strengths. ....   | 138 |
| 7.2.1 The $R/\eta$ method. ....   | 139 |
| 7.2.2 The $2R_2-R_1$ method. ....   | 140 |
| 7.2.3 The $2\eta_{xy} - \eta_z$ method.....   | 141 |
| 7.2.4 Robust Analysis of Data. ....   | 142 |
| 7.3 Analyses of Relaxation Data Using the Lipari-Szabo Approximation. ....  | 144 |
| 7.3.1 “Standard” Lipari-Szabo approach (LS). ....   | 144 |
| 7.3.2 Lipari-Szabo Approach Including the CSA (LS-CSA). ....  | 145 |
| 7.3.3 Lipari-Szabo Analysis of Spectral Densities (LS-SDF). ....  | 146 |
| 7.4 Site-Specific CSAs in GB3.....  | 147 |

|  |     |
|--|-----|
| 7.4.1 Site Specific $\Delta\sigma$ , $\Delta\sigma_g$ , and the Principal Values and Orientations of the Assumed Axially Symmetric Chemical Shielding Tensors.....           | 147 |
| 7.4.2 Correlation of $\Delta\sigma$ with Isotropic Chemical Shift and Structure. ....  | 158 |
| 7.4.5 Variability in the CSAs in GB3 Separated from Uncertainty.....   | 162 |
| 7.5 Backbone Order Parameters in GB3 from Multifield $^{15}\text{N}$ Relaxation Data; The Effect of Site-Specific versus Uniform $^{15}\text{N}$ CSAs.....                   | 169 |
| 7.5.1 Backbone Order Parameters: Assuming a Uniform $^{15}\text{N}$ CSA. ....  | 169 |
| 7.5.2 Backbone Order Parameters: the Effect of Site-Specific $^{15}\text{N}$ CSAs.....   | 174 |
| 7.5.3 LS Fit of the Spectral Densities Directly.....   | 177 |
| 7.6 Discussion and Conclusions .....   | 178 |
| 7.6.1 Possible Sources of Systematic Errors in $^{15}\text{N}$ CSA Determination from Multiple-Field Data.....   | 178 |
| 7.6.2 Conclusions.....   | 185 |
| Chapter 8: Novel Solution-State Relaxation Measurements .....  | 187 |
| 8.1 Motivation.....  | 187 |
| 8.2 What is the Effect of $^1\text{H} \rightarrow ^2\text{H}$ Isotopic Substitution on the CSA of the $^{15}\text{N}$ Nucleus in a Peptide? A Quantum Mechanical Study ..... | 193 |
| 8.2.1 Expectation Based on the Observed Effect on the $^{15}\text{N}$ Isotropic Chemical Shift.....  | 193 |
| 8.2.2 <i>Ab Initio</i> Quantum Mechanical Calculations.....  | 196 |
| 8.2.3 Conclusions.....   | 200 |
| 8.3 Pulse Sequences, Spectra, and Rates .....  | 201 |
| 8.3.1 Pulse Sequences for Measurement of $^{15}\text{N}$ $R_1$ and $^{15}\text{N}$ $R_2$ using Direct Nitrogen Detection.....  | 201 |
| 8.3.2 Measured Rates.....  | 204 |
| 8.4 Diffusion Tensor of GB3 in $\text{D}_2\text{O}$ Solution Derived from $^{15}\text{N}$ Direct Detection Relaxation Rates: Proof of Principle.....                         | 206 |
| 8.4.1 Subtraction of Contributions to the Relaxation Rates Based On Model-Free Parameterization .....  | 206 |
| 8.4.2 Diffusion Tensor of GB3: Proof of Principle .....  | 208 |
| 8.4.3 Relative Orientation of the CSA and Dipolar Relaxation Mechanisms ..   | 211 |
| 8.5 Conclusions.....   | 213 |
| Chapter 9: Summary and Concluding Remarks.....   | 215 |
| 9.1 Summary of Results.....  | 215 |
| 9.2 Scope for Future Studies.....  | 220 |
| Appendix A.....  | 222 |

## List of Tables

- 1.3.1  $^{15}\text{N}$  chemical shift tensor principal values from solid-state NMR measurements
- 2.2.1 The Dipole-Dipole and CSA Hamiltonians in the laboratory frame
- 3.2.1 Relaxation delays for all experiments
- 5.3.1 Hydrodynamic characteristics of the GB3 domain derived from  $^{15}\text{N}$  relaxation data using various models of the overall tumbling and from hydrodynamics calculations
- 7.1.1 Characteristics of the overall rotational diffusion tensor of GB3 derived from  $^{15}\text{N}$  relaxation data at five magnetic fields
- 7.1.2 Characteristics of the overall rotational diffusion tensor derived from  $^{15}\text{N}$  relaxation data at five magnetic fields using RDC refined structures of GB3
- 7.4.1 Statistics of the  $^{15}\text{N}$  CSA values in GB3 determined here using several methods.
- 7.4.2 Site-specific components of the chemical shift and shielding tensors (assumed axially symmetric)
- 7.6.1 Characteristics of the fully anisotropic overall rotational diffusion tensor of GB3 derived from  $^{15}\text{N}$  relaxation data at five magnetic fields
- 8.1.1 Percent contribution to  $^{15}\text{N}$   $R_1$  and  $R_2$  in  $^{15}\text{N}$ - $\{^1\text{H}\}$  and  $^{15}\text{N}$ - $\{^2\text{H}\}$  spin systems at 800 MHz from dipolar and CSA relaxation mechanisms for a protein with a tumbling time of 5 ns
- 8.1.2 Calculated estimates of the contributions to  $^{15}\text{N}$   $R_1$  and  $R_2$  in  $^{15}\text{N}$ - $\{^1\text{H}\}$  and  $^{15}\text{N}$ - $\{^2\text{H}\}$  spin systems in GB3 at five fields from various relaxation mechanisms

8.4.1 Characteristics of the overall rotational diffusion tensor of GB3 determined from  $^{15}\text{N}$  relaxation data measured by the direct method and by conventional relaxation experiments



## List of Figures

- 1.2.1 Schematic illustration of polypeptide chain
- 3.1.1 Pulse sequence for decoupled heteronuclear single quantum coherence (HSQC) experiment
- 3.1.2 HSQC spectrum of the GB3 domain
- 3.1.3 Pulse sequence for measurement of  $^{15}\text{N}$  longitudinal relaxation rate  $R_1$
- 3.1.4 Pulse sequence for measurement of  $^{15}\text{N}$  transverse relaxation rate  $R_2$
- 3.1.5 Pulse sequence for measurement of the steady-state heteronuclear  $^{15}\text{N}$ - $\{^1\text{H}\}$  nuclear Overhauser effect (NOE).
- 3.2.1 Linear fit of the signal intensity decay for Tyr3 residue in  $R_1$  and  $R_2$  relaxation experiments at 14.1 T
- 4.1.1 Schematic illustration of  $^{15}\text{N}$  CSA/dipole-dipole cross-correlation in a 2D model molecule
- 4.2.1 Pulse sequence for the measurement of transverse  $^{15}\text{N}$  CSA/dipole-dipole CCR from a coupled HSQC spectrum—in-phase (IP) experiment
- 4.2.2 Pulse sequence for the measurement of transverse  $^{15}\text{N}$  CSA/dipole-dipole CCR from a coupled HSQC spectrum—anti-phase (AP) experiment
- 4.2.3 Sections of spectra recorded using the IP and AP pulse sequences and sections of the addition and subtraction simplified spectra
- 4.2.4 Representative decay curves for the ratio of peak volumes and peak intensities from the IP and IPAP experiments
- 4.2.5 The agreement between  $\eta_{xy}$  measured using the IP and IPAP methods with the A/B method

4.2.6 Pulse sequence for measurement of transverse  $^{15}\text{N}$  CSA/dipole-dipole CCR using conventional (not constant time)  $t_1$  evolution, IP and AP experiments

4.2.7 The agreement between  $\eta_{xy}$  values measured using the conventional  $t_1$  evolution IPAP experiment and the constant time  $t_1$  evolution IPAP experiment

4.3.1 The  $\text{S}^3\text{E}$  selective pulse sequence for measurement of the transverse  $^{15}\text{N}$  CSA/dipole-dipole CCR using conventional and constant time  $t_1$  evolution

4.3.2 Correlation between  $\eta_{xy}$  measured at 14.1 T using the IP and IPAP methods with  $\eta_{xy}$  measured using the  $\text{S}^3\text{E}$  method

4.4.1 Pulse sequence for measurement of longitudinal  $^{15}\text{N}$  CSA/dipole-dipole CCR

4.4.2 Agreement of longitudinal  $^{15}\text{N}$  CSA/dipole-dipole CCRs in protonated and deuterated GB3 samples measured using IPAP experiment

5.3.1 Relaxation rates in GB3 at 14.1 T

5.3.2 Ribbon representation of tertiary structure of GB3 domain with the orientation of the diffusion tensor axes indicated by rods

5.3.3 Orientation dependence of  $\rho^{\text{exp}}$

5.3.4  $\rho^{\text{exp}}$  and  $\rho^{\text{calc}}$  for axially symmetric and fully anisotropic diffusion tensor

5.4.1 Comparison of experimental  $\tau_c$  and  $D_{\parallel} / D_{\perp}$  with results of theoretical hydrodynamic bead model

5.4.2 Comparison of experimental  $\tau_c$  and  $D_{\parallel} / D_{\perp}$  with results of theoretical hydrodynamic shell model

6.2.1 Linear agreement between experimental values of  $\eta_{xy}$  and  $R_2'$  at 14.1 T

6.2.2 Comparison of the measured values of  $R_2'$  with their “exchange-free” estimates  $R_{2\text{free}}'$

6.3.1 Comparison of the model-free parameters for backbone dynamics in GB3, determined using the three models of the overall tumbling

6.4.1  $R_{\text{ex}}$  values calculated from the experimental data at 9.4 and 14.1 Tesla assuming -160 ppm as an average  $^{15}\text{N}$  CSA value.

6.5.1  $R_{\text{ex}}$  contribution to  $R_2$  based on the first 12 NMR structures from the 24 structure bundle (2IGH.pdb) assuming a fully anisotropic diffusion tensor.

6.6.1 Conformational exchange in GB3, determined using the relaxation compensated CPMG sequence.

7.1.1 The agreement between the measured  $R_2$ s and their reconstructed “exchange-free” values,  $R'_{2\text{free}} = R_1' \cdot \eta_{\text{xy}} / \eta_z$  at multiple field strengths

7.2.1 Representative fits of the dependence of  $2R_2' - R_1'$  on  $\omega_N^2$ .

7.4.1 The site-specific  $^{15}\text{N}$  CSAs, from the  $2R_2 - R_1$  method,  $R/\eta$  method, and the LS-CSA method

7.4.2 The values of  $\Delta\sigma_g$  and the  $\beta_z$  angles from the  $R/\eta$  and  $2\eta_{\text{xy}} - \eta_z$  methods

7.4.3 Site-specific  $^{15}\text{N}$  CSA values, averaged over all three methods, show significant CSA variability in GB3

7.4.4 CSAs in GB3 as a function of their isotropic chemical shifts

7.4.5 CSAs in GB3 as a function of amino acid type.

7.4.6 Residue-specific values of the anisotropy of the chemical shielding tensor,  $\Delta\sigma$ , normalized by the isotropic chemical shift and as a function of the backbone dihedral angles,  $\phi$ ,  $\psi$ , and  $\omega$

7.4.7 The likelihood functions obtained from different methods illustrate the significant site-so-site variability in the  $^{15}\text{N}$  CSA values.

7.5.1 Backbone order parameters determined from  $^{15}\text{N}$  relaxation data at each field using different CSA models.

7.5.2 The agreement between the spectral density component,  $J(0)$ , measured using the  $2R_2-R_1$  method and reconstructed from the LS parameters.

7.5.3 Representative LS fit of all spectral density components from the five-field measurements for Phe30.

7.6.1 Comparison of CSAs in GB3 derived using the LS-CSA method with fully anisotropic and axially symmetric rotational diffusion tensors.

8.2.1 Geometry of the N-methylacetamide (NMA) molecule

8.2.2 Schematic illustration of the expected orientation of the principal components (eigenvectors) of the  $^{15}\text{N}$  CSA tensor with respect to the peptide plane.

8.3.1 Pulse sequences for 1D  $^{15}\text{N}$ -detected  $R_1$  measurement.

8.3.2 Pulse sequences for 1D  $^{15}\text{N}$ -detected  $R_2$  measurement.

8.3.3  $^{15}\text{N}$ -detected spectrum of GB3 (pH 5.5, 24°C), recorded at 11.7 T.

8.3.4 Representative fitting curves for  $R_1$  and  $R_2$  measurements for Glu24 and Lys50.

8.3.5. Longitudinal  $R_1$  and transverse  $R_2$   $^{15}\text{N}$  relaxation rates in GB3 in  $\text{D}_2\text{O}$ , determined using  $^{15}\text{N}$  direct detection at 11.5 T.

8.4.1. Variation in the residuals of the fit (from ROTDIF analysis), as the main interaction frame is tilted away from the direction of the NH vector, while remaining in the peptide plane (angle  $\xi = 0$ ).

# Chapter 1: Introduction and Specific Aims

## **1. 1 Introduction**

An appreciation of biomolecular flexibility and dynamics is essential to an understanding of the processes necessary to life. The many roles of proteins in these processes (proteins make up structural elements of cells and organisms, enzymes catalyze biochemical reactions, transport proteins such as hemoglobin facilitate movement of essential chemical compounds, and generally proteins are involved in myriad biological processes) ultimately reflects both the large amount of information encoded in protein structures and the large range over which this information content can be varied by modification of these structures (dynamics). A key characteristic of proteins is their functional specificity—for example, a particular enzyme will bind a specific substrate partner and catalyze a specific chemical reaction. Frequently this specificity is such that a small change in the structure or dynamics of the enzyme or the ligand can dramatically change their binding affinity, and the activities of many enzymes are regulated by structural and dynamic changes induced by interactions with other molecules<sup>1,2</sup>. In order to comprehend the ways in which proteins achieve their biological functionality, we have to understand the ways in which the many parts of a protein molecule move and interact and to what extent these various motions are determined by molecular structure and intermolecular forces. Understanding the processes underlying protein stability, recognition, specificity, and catalysis requires knowledge of the atomic-level details of protein structure and dynamics.

Nuclear magnetic resonance (NMR) is conceptually beautiful and extraordinarily useful. NMR has been used to demonstrate textbook-like examples of the most fundamental principles of quantum mechanics<sup>3,4</sup>. NMR can be used to determine four-dimensional (three spatial coordinates plus time) structures of molecules with intermolecular distance accuracy that “seemingly” violates the theoretical limit given by Heisenberg’s uncertainty relation<sup>5</sup>. NMR is used daily in hospitals to non-invasively image human tissue for disease diagnosis. NMR provided the first experimental realization of a quantum computer<sup>6,7</sup>.

For biology, NMR has been used to solve the structures of more than 5,000 proteins and nucleic acids. Biomolecular structure determination (by NMR, X-ray crystallography, and other techniques) has revolutionized biochemistry. The uniqueness of NMR in this realm, however, is not in its ability to report on structure, currently it’s being outdistanced by crystallography in terms of resolution and output (X-ray crystal structures make up ~86% of the protein structures deposited in the protein data bank, while NMR structures make up only ~14%), but in its ability to report on biomolecules in solution. This is an obvious advantage, since it is in solution where most biomolecules perform their physiological function, but since molecules in solution are constantly tumbling and moving with thermal energy, solution NMR of biomolecules must necessarily include consideration of motions in the systems under study. NMR spin relaxation is caused by molecular motion<sup>8</sup>, and spin relaxation determines the linewidths and thus the intensities of NMR signals. An NMR spectroscopist who studies proteins, therefore, even if their purpose is structure determination, must consider relaxation effects, and thus molecular motion, to

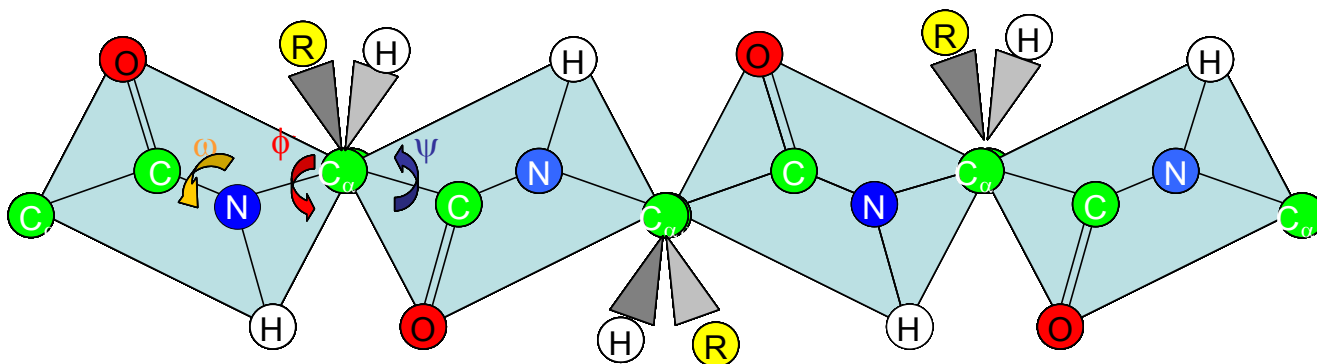
interpret spectra and design experiments. Because of the connection between structure and dynamics in NMR spectra, one possible future direction for NMR spectroscopy of biomolecules is simultaneous structure and dynamics determination<sup>9</sup>.

## **1.2 Motion in Proteins**

A brief review of protein structural and energetic properties is necessary as an introduction to consideration of their dynamics. A protein is essentially a large number of chemical groups linked by covalent bonds. A short length of polypeptide chain is shown in Fig 1.2.1. The atoms indicated in cyan make up what is termed the peptide plane. The partial double bond character of the CONH bond causes it to be impermissive to rotation, so the  $\omega$  torsion angle typically has values very close to  $180^\circ$  (or  $0^\circ$  for the less common cys stereoisomer) in proteins. However, many of the other covalent bonds are rotationally permissive, which allows the protein backbone some degree of flexibility. Steric repulsion is the primary limitation to the ranges of rotation of the torsion (dihedral) angles,  $\phi$  and  $\psi$ , of a polypeptide. Librational movement of individual bonds (such as the backbone amide bond) are expected to be fairly uniform in magnitude along the polypeptide backbone<sup>10,11</sup>, and backbone conformational fluctuations of  $\phi$  and  $\psi$  angles on the sub-nanosecond timescale are by-nature anisotropic (with director postulated to be along the  $C_\alpha$ - $C_\alpha$  axis). When a protein is folded into a globular structure consisting primarily of  $\alpha$ -helices and  $\beta$ -strands, the steric hindrance of the folded conformation causes the torsion angles in the secondary structure elements to take on characteristic values (visible from a Ramachandran plot) and affects the energy potential which determines the

fluctuations about these values. The steric hindrance of the particular type of amino acid side-chain group will also affect this potential as will any intramolecular interactions. Therefore the amplitudes of the anisotropic motions of a polypeptide chain will depend on its secondary and tertiary structure and show variation from residue to residue.

**Figure 1.2.1. Schematic of a short length of polypeptide chain in a fully-extended conformation. Here R represents the side chain group and can be any of 19 of the 20 naturally occurring amino acids (Proline is predominately in the trans stereoisomer, not illustrated here). The  $\pi$  orbitals of the partial double bond between C'-O and N-C' creates the plane like character of the atoms indicated by the rectangles. The dihedral angles  $\psi$  and  $\phi$ , through which the peptide planes can rotate, are indicated by arrows near the first alpha carbon.**



Motions in proteins span a broad range of amplitudes and timescales. For example, a recent review of oxygen binding/release motions in myoglobin<sup>12</sup> cited examples of dynamic events ranging from femtosecond timescales all the way to the millisecond timescale. Protein motions can be broadly categorized as motions which are fast and highly localized, such as atomic level fluctuations (bond stretching and bond librations occur on fs-ps timescales), intermediate timescale motions which involve the collective motion of several atoms (unhindered surface sidechain motions occur on ps-ns timescales, “breathing” motions of the protein backbone occur on a ns timescale) and motions which occur on the scale of the whole molecule and over



longer times (the so-called conformational exchange motions—collective motion of large groups of atoms—occur on  $\mu$ s-ms timescales). However, the various motions in a protein are coupled to one another, and large-amplitude/long-timescale dynamic transitions naturally involve many smaller intermediate-amplitude/intermediate-timescale transitions which involve many fast, local, atomic-level motions as well.

There are several compelling arguments for the biological relevance of fast motions in the peptide planes and side chains. Besides several examples of fast motions during high energy events in proteins (e.g. bond breaking)<sup>12,13</sup>, the kinetic energy in the fast, low amplitude, high probability fluctuations in backbone  $\phi/\psi$  torsion angles is a source of energy for large scale conformational changes<sup>14</sup>. Larger-than-average-amplitude motions on the ps-ns timescale can therefore be suggestive of motions which occur during ligand binding as the fast motions indicate flexibility<sup>14</sup>. Furthermore, motions do not have to occur on the same time scale as a particular biological reaction in order to influence the reaction. When considered thermodynamically, the free energy of any reaction is a balance of enthalpic and entropic contributions to the free-energies of the initial and final states. Entropic effects due to changes in fast internal dynamics associated with molecular recognition processes have been shown to have a profound impact on binding affinities<sup>2,15</sup>, though the binding itself often takes place on a much slower ( $\mu$ s-ms) timescale.

Additionally, the motions occurring in proteins while free in solution have relevance to their biological function. The postulation (originally advanced by Frauenfelder<sup>16</sup>) that the “information content” for all motions (such as

conformational exchange motion upon ligand binding or the folding transition) is contained within (and therefore potentially extractable from) the thermally driven motions of the polypeptide backbone is recently being reasserted with growing credibility. Eisenmesser et al. have shown that the exchange motions that occur in the enzyme cyclophilin A during its function as a prolyl cis-trans isomerase are already present in the free enzyme <sup>17,18</sup>. This implies that some functional motions in enzymes may be due to an “intrinsic plasticity”, determined by the enzyme’s amino-acid sequence. Intrinsic flexibility in biomolecules is also reflected in recent theoretical work <sup>19</sup>, which uses a simple harmonic oscillator based potential applied to protein structures to successfully predict large-scale, low frequency normal-mode motions in proteins <sup>20</sup>. Motions predicted in such a fashion were found to have good agreement with experimentally determined motions observed during protein-ligand binding, allostery, and catalysis <sup>20-23</sup>, these motions are likely, therefore, to be “intrinsic”, i.e. determined by the structure and coded for by the amino acid sequence of the protein, and therefore under control of genetic selection.

Techniques for the study of protein dynamics range from theoretical and computational modeling to measurement by X-ray diffraction, neutron and optical scattering, perturbation-relaxation type experiments (such as temperature- or pH-jump experiments), time-resolved fluorescence spectroscopy, atomic force microscopy, and solution and solid-state NMR. Each dynamic process (i.e. motion) has a characteristic amplitude, time-scale, and energy and certain techniques are well suited for the study of specific types and timescales of protein motions.

One of the most effective and widely used techniques for the study of internal protein dynamics on ps-ns timescales is solution NMR spin relaxation. In kinetics “relaxation” refers to the return of a system to equilibrium after a perturbation. NMR spin relaxation experiments are unique among perturbation-relaxation type experiments in that the perturbed variables (the nuclear spin degrees of freedom) involve energies which are very small ( $\hbar\omega_H$  at 600 MHz is  $< 4 \cdot 10^{-25}$  Joules, which is  $\sim 10,000$  times smaller than  $k_B T$  at  $20^\circ\text{C}$ ) and very weakly coupled to the ordinary conformational dynamics of the molecule. This means that the ordinary chemical energies and dynamic properties of the system are almost entirely unaffected by spin-flips caused by NMR pulses. NMR spectroscopy also enables investigation of motions at many atomic sites simultaneously, while methods such as stopped-flow fluorescence, fluorescent resonance energy transfer, and atomic force microscopy report on global motions of the molecule or dynamics at a particular reporter site. Furthermore, the connections between the coherences which give rise to peaks in an NMR spectrum and structural aspects of proteins is well established due to the extensive use of NMR in protein structure determination. Much insight into protein dynamics has been gained by NMR relaxation experiments.

There are a multitude of NMR experiments for studying motion in proteins. In particular,  $^{15}\text{N}$  auto-relaxation, cross-relaxation, and cross-correlation experiments (e.g.  $^{15}\text{N}$   $R_1$  and  $R_2$ ,  $^{15}\text{N}\{^1\text{H}\}$  NOE, and cross-correlation rates) have been shown to give useful information about backbone motions on the ps-ns timescale.  $R_2$  is also, in principle, sensitive to motions on slower timescales. These experiments have their origins along with the very first NMR experiments<sup>24, 8</sup> and development of new

techniques and improvements to existing sequences are still underway (e.g. <sup>25,26</sup>).

There are a variety of methods for the derivation of motional parameters from spin-relaxation data. Among these, the most frequently employed are the model-free analysis of Lipari and Szabo<sup>27,28</sup>, the extended model-free analysis proposed by Clore et al. <sup>29</sup>, and spectral density function mapping <sup>30,31</sup>. Model-free analyses involve fitting relaxation rates to a specific form of the spectral density function (a sum of Lorentzian functions) involving parameters which describe the amplitude and timescale of the motions of the backbone bonds. The model-free parameters are extracted by minimization of a target function. It is termed “model-free” since it involves no assumptions about the model of motion (wobbling in a cone, Gaussian-angular-fluctuations about specific orientations, jumping from site-to-site, etc. are examples of specific models of motion) or the specific trajectory of the bond vector, though it does assume a specific approximate form of the correlation/spectral density function of the motion (see Chapter 2). Model-free parameters have been reported for around 100 proteins, and have given insight into the dynamics involved in several important biological processes by providing a picture of local motions.

Lipari-Szabo model-free parameters from <sup>15</sup>N relaxation rates report on the timescale and degree of spatial restriction of the backbone N-H bond. These parameters reflect all motions of this bond on the ps-ns timescale—including very fast vibrational motion and fast bond librations<sup>27,28</sup>, though it has been shown that the largest contribution to this motion comes from the motions of the peptide planes and larger-scale segmental motions <sup>32,11</sup>. In principle, these parameters contain a wealth of information about the individual or concerted motions of all the peptide planes in a

protein and can be used to test different motional models, to estimate conformational entropy, to validate molecular dynamics simulations, etc. However, in practice, there are many known sources of assumptions/errors/ambiguities in fitting relaxation data to an analytical form of the spectral density function, and these errors frequently preclude quantitative analysis of model-free parameters.

$^{15}\text{N}$  spin relaxation rates are sensitive to any motion that modulates the effective magnetic field at the location of the  $^{15}\text{N}$  nucleus. This includes all reorientations with respect to the external magnetic field, any motion that modulates the through-space dipolar interaction with other nuclei, and any motion which causes interference of this dipolar interaction with other relaxation mechanisms (see Chapter 4)<sup>33</sup>. Deconvolution of the effects of these motions into conceptually tractable parts (e.g. conformational exchange, overall tumbling, local motion of peptide planes, fast bond librations) can lead to uncertainty in motional parameters. Overall rotational diffusion is generally a larger cause of spin relaxation than fast, local motion of individual bonds or collective motions of groups of atoms<sup>34</sup>, therefore the correct deconvolution of the overall tumbling is essential to an accurate picture of local motion.

An additional source of inaccuracy of NMR relaxation derived motional parameters comes from inaccurate estimation of interaction parameters that influence relaxation. The values of these parameters are not known with adequate precision for relaxation analysis, and variation from residue-to-residue of these interaction parameters is rarely taken into account. This will be discussed in the following section.

### **1.3 Chemical Shielding Anisotropies in Proteins and Peptides**

#### **1.3.1 Chemical Shieldings and their Relationship to Protein Structure**

Variation in the chemical shielding of nuclei makes NMR spectroscopy of complex biomolecules like proteins possible. It is this variation which causes magnetically inequivalent nuclei in a molecule to be distinguishable in an NMR spectrum. The field felt by any given nucleus is slightly different from the applied magnetic field due to charged electrons circulating the nucleus. This difference in the field felt by a particular nucleus in a molecule and that which would be felt by a theoretical bare nucleus (i.e. the applied external field) is called the chemical shielding of the nucleus. This quantity can be described as a fraction of the applied magnetic field, varies from site to site within the molecule, and contains a wealth of information for molecular structure prediction. The chemical shielding, represented by  $\sigma$ , is observed through the NMR chemical shift,  $\delta$ , which is related to the chemical shielding by the relationship:

$$\delta = \frac{(\sigma_{ref} - \sigma)}{1 - \sigma_{ref}} \approx \sigma_{ref} - \sigma, \quad (1.3.1)$$

where  $\sigma_{ref}$  is the chemical shielding of the same nucleus in a reference compound.

NMR chemical shifts are extremely sensitive to the chemical environment of nuclei and change with temperature, pH, molecular geometry, and in the presence of intermolecular interactions. The strong dependence of isotropic chemical shifts on protein structure has long been recognized—in particular, the correlation between  $^1\text{H}^\alpha$  chemical shift and secondary structure has been studied, the periodicity of the  $\text{H}^\text{N}$  chemical shift has been shown to reflect hydrogen bond length<sup>35</sup>, and the secondary

$^{13}\text{C}^\alpha$  and  $^{13}\text{C}^\beta$  chemical shifts have been shown to correlate with dihedral angles<sup>36,37</sup>.

Computer programs have been developed that obtain backbone torsion angle restraints from  $^1\text{H}^\alpha$  and  $^{13}\text{C}^\alpha$ , and the very useful algorithm TALOS by Cornilescu et al.<sup>38</sup> searches an empirical database for both sequence and chemical shift homology, and uses this information to predict backbone angles for protein structure calculation. Changes in  $^{15}\text{N}$  and  $^1\text{H}$  isotropic chemical shifts derived from conformational exchange contributions to relaxation rates have even been used to solve the structures of “invisible”<sup>39</sup> extremely low-population molecular states<sup>40</sup>.

### 1.3.2 The Chemical Shielding Tensor

The above discussion concerns what is termed the isotropic chemical shift. In most solution NMR experiments, the motion of the molecule averages the chemical shielding tensor of the nuclei under study so that only the isotropic value of the chemical shift remains. However, the local electronic environment of a nucleus is almost never isotropic, and the anisotropy of the chemical shielding tensor is an important source of nuclear spin relaxation. The expression for this chemical shielding anisotropy (CSA) in terms of the principal components (eigenvalues) of the chemical shielding tensor is:

$$|\Delta\sigma| = [\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 - (\sigma_{xx}\sigma_{yy} + \sigma_{xx}\sigma_{zz} + \sigma_{yy}\sigma_{zz})]^{1/2} \quad (1.3.2)$$

### 1.3.3 Chemical Shielding Tensors in Proteins and Peptides

Much is known about chemical shift tensors in molecules.  $^{13}\text{C}$ ,  $^1\text{H}$ ,  $^{19}\text{F}$ ,  $^{32}\text{P}$  and  $^{15}\text{N}$  shielding tensors have been measured from solid state NMR and theoretically determined from quantum chemical calculations.

There is considerable debate surrounding both the variability and the average values of the  $^{15}\text{N}$  chemical shielding anisotropy in proteins. Previous to the 1990's, most experimental information on the  $^{15}\text{N}$  CSA had come from solid-state NMR powder patterns or from rotation studies by solid-state NMR on crystals of small peptides (although only one CSA tensor has so far been reported from solid-state NMR on a single crystal <sup>41</sup>). These studies provide a relatively narrow distribution of CSA's, with a mean value of -156.0 ppm and a standard deviation of  $\sim 5.7$  ppm <sup>34</sup>. Table 1.3.1 shows a selection of solid state NMR measurements of  $^{15}\text{N}$  chemical shift tensors in peptides. The information in this table is taken from <sup>34</sup>.



**Table 1.3.1**  $^{15}\text{N}$  chemical shift tensor principal values from solid-state NMR measurements. The information for this table was taken from Korzhnev (Korzhnev et al., *NMR studies of Brownian tumbling and internal motions in proteins*, Prog. in NMR Spect. 2001). See legend below for more information

| Sample  | $\delta_{xx}^a$ | $\delta_{yy}^a$ | $\delta_{zz}^a$ | $\delta_{iso}^b$ | $\beta^c$  | $\Delta\delta = -\Delta\sigma^d$ |
|---|-----------------|-----------------|-----------------|------------------|------------|----------------------------------|
| Ala*Ala   | 65.3            | 78.1            | 215.5           | 119.6            | 12.6°      | -144                             |
| AcGly*AlaNH <sub>2</sub>                            | 44.6            | 85.1            | 229.4           | 119.7            | 17.6° ± 2° | -164.55                          |
| (*Ala) <sub>n</sub> α-helix                         | 47.7            | 64.1            | 213.7           | 108.5            |            | -157.8                           |
| (*Ala) <sub>n</sub> -5 β-sheet                      | 53.7            | 71.4            | 210.7           | 111.9            |            | -148.2                           |
| (*Ala,Leu) <sub>n</sub> α-helix                     | 44.7            | 66.6            | 213.7           | 108.3            |            | -158.1                           |
| (*Ala,Asp (OBzl)) <sub>n</sub> α-helix              | 47.7            | 68.4            | 217.7           | 111.3            |            | -159.7                           |
| (*Ala,Glu (OBzl)) <sub>n</sub> α-helix              | 48.7            | 66.4            | 215.7           | 110.3            |            | -158.2                           |
| (*Ala,Glu (Ome)) <sub>n</sub> α-helix               | 46.7            | 67.8            | 214.7           | 109.7            |            | -157.5                           |
| (*Ala, Val) <sub>n</sub> β-sheet                    | 44.7            | 72.1            | 211.7           | 109.5            |            | -153.3                           |
| (*Ala, Ile) <sub>n</sub> β-sheet                    | 49.7            | 72.7            | 209.7           | 110.7            |            | -148.5                           |
| (*Asp (OBzl)) <sub>n</sub> -1 α <sub>R</sub> -helix | 48.7            | 62.5            | 214.7           | 108.6            |            | -159.1                           |
| (*Asp (OBzl)) <sub>n</sub> -2 α <sub>L</sub> -helix | 50.7            | 58              | 210.7           | 106.5            |            | -156.35                          |
| (*Asp (OBzl)) <sub>n</sub> -2 α <sub>L</sub> -helix | 49.7            | 57.1            | 211.7           | 106              |            | -158.3                           |
| (*Asp (OBzl)) <sub>n</sub> -2 α <sub>L</sub> -helix | 50.7            | 66.1            | 212.7           | 109.7            |            | -154.3                           |
| N-Ac*Gly  | 37.0            | 82.8            | 220.4           | 113.4            | 25.5° ± 1° | -160.50                          |
| (*Gly) collagen powder                              | 42.3            | 67.0            | 223.4           | 110.9            | 24.5° ± 1° | -168.75                          |
| (*Gly) collagen oriented                            | 42.3            | 67.0            | 223.4           | 110.9            | 24.5° ± 2° | -168.75                          |
| (*Gly) collagen                                     | 45.6            | 67.6            | 216.8           | 110.0            | 23°        | -160.20                          |
| (*Gly) maganine                                     | 42.0            | 73.2            | 215.0           | 110.1            | 22° ± 2°   | -157.40                          |
| Boc-(Gly) <sub>2</sub> *Gly-OBzl                    | 55.1            | 62.1            | 223.0           | 113.4            | 22° ± 1°   | -164.40                          |
| Boc-(Gly) <sub>2</sub> *Gly-OBzl                    | 36.4            | 83.4            | 220.4           | 113.4            | 24° ± 1°   | -160.50                          |
| Gly*Gly   | 46.8            | 79.7            | 220.8           | 115.8            |            | -157.55                          |
| Gly *Gly·HCL  | 57.3            | 59.8            | 210.0           | 109.0            | 18.6° ± 2° | -151.45                          |
| AcGly *GlyNH <sub>2</sub>                           | 40.7            | 64.2            | 210.0           | 105.0            | 17.6° ± 2° | -157.55                          |
| Gly *Gly·HCL·H <sub>2</sub> O (powder)              | 58.5            | 64.1            | 209.5           | 110.7            | 25° ± 5°   | -148.20                          |
| Gly *Gly·HCL·H <sub>2</sub> O (crystal)             | 60.3            | 70.9            | 215.9           | 115.7            | 21.3°      | -150.30                          |
| (*Gly) <sub>n</sub> β-sheet                         | 45.7            | 61.4            | 205.7           | 104.3            |            | -152.15                          |
| (*Gly) <sub>n</sub> 3 <sub>10</sub> -helix          | 49.7            | 62.8            | 214.7           | 109.1            |            | -158.45                          |
| (*Gly, Ala) <sub>n</sub> α-helix                    | 44.7            | 57.6            | 212.7           | 105.0            |            | -161.55                          |
| (*Gly,Ala) <sub>n</sub> β-sheet                     | 39.7            | 66.0            | 206.7           | 104.1            |            | -153.85                          |
| (*Gly,Leu) <sub>n</sub> α-helix                     | 45.7            | 61.7            | 210.7           | 106.0            |            | -157.00                          |
| (*Gly,Leu) <sub>n</sub> β-sheet                     | 40.7            | 66.2            | 206.7           | 104.5            |            | -153.25                          |
| (*Gly, Val) <sub>n</sub> β-sheet                    | 39.7            | 74.6            | 203.7           | 106.0            |            | -146.55                          |
| (*Gly, Ile) <sub>n</sub> β-sheet                    | 45.8            | 68.3            | 209.7           | 108.6            |            | -152.65                          |
| (*Gly,Lys(Z)) <sub>n</sub> α-helix                  | 40.7            | 69.2            | 208.7           | 106.2            |            | -153.75                          |
| (*Gly,Glu(OBzl)) <sub>n</sub> α-helix               | 47.7            | 61.2            | 210.7           | 106.5            |            | -156.25                          |
| (*Gly,Sar) <sub>n</sub>                             | 38.7            | 65.8            | 204.7           | 103.1            |            | -152.45                          |
| (*Phe) maganine                                     | 55.0            | 80.0            | 220.0           | 118.3            | 22° ± 3°   | -152.5                           |
| AcGly *TyrNH <sub>2</sub>                           | 52.1            | 77.1            | 209.3           | 112.8            | 19.6° ± 2° | -144.70                          |

<sup>a</sup> Here  $\delta_{zz} > \delta_{yy} > \delta_{xx}$  are the frequency ordered principal values of the chemical shift tensor, and  $\sigma_{zz} < \sigma_{yy} < \sigma_{xx}$  (i.e  $\sigma_{zz}$  is the least shielded component) are the principal values of the corresponding shielding tensor. The relationship between these two conventions for the tensor is  $\underline{\sigma} = -(\underline{\delta} - \underline{1} \cdot \delta^{iso})$  where  $\underline{\delta}$  is the chemical shift tensor and  $\underline{\sigma}$  is the chemical shielding tensor. <sup>b</sup> The isotropic chemical shift  $\delta_{iso} = (\delta_{xx} + \delta_{yy} + \delta_{zz})/3$ . <sup>c</sup> The angle between the least shielded axis,  $\delta_{zz}$ , of the CSA tensor and the N-H bond. <sup>d</sup>  $\Delta\delta$  is the chemical shift anisotropy and is defined as  $\Delta\delta = (\delta_{xx} + \delta_{yy})/2 - \delta_{zz}$ . The chemical shielding anisotropy is therefore given by  $\Delta\sigma = -\Delta\delta = \sigma_{zz} - (\sigma_{xx} + \sigma_{yy})/2$ .

Based on these type of measurements, early relaxation studies of backbone dynamics in proteins commonly assumed a value for the  $^{15}\text{N}$  CSA of -160 ppm in conjunction with a NH bond length of  $1.02 \text{ \AA}$ <sup>42</sup>, though it was noted<sup>43,44</sup> that a value of -170 ppm provided a better fit to relaxation data in protein systems. CSA values can in principle be measured from changes in chemical shifts induced by weak alignment in a magnetic field, but the only studies thus far have not been able to distinguish site-specific CSA's due to limited precision of available data, but have reported an average CSA value of -173 ppm<sup>45</sup> in the protein Ubiquitin and between -172.1 and -174.4 in hen Lysozyme<sup>46</sup>.

Quantum mechanical calculations, however, indicate that the amide  $^{15}\text{N}$  CSA tensors in a protein depend on several parameters, including local backbone geometry and hydrogen bonding, which may vary substantially from residue to residue within a protein<sup>47 48</sup>. Stretching of the H-N bond, caused by hydrogen bonding in a peptide has previously been studied by *ab initio* calculation. These effects on the CSA are small,  $\sim 3\text{-}9 \text{ ppm}$ <sup>47,49</sup>, when compared with the dependence of the  $^{15}\text{N}$  CSA values on backbone conformation from a similar calculation, which was  $\sim 30\text{-}40 \text{ ppm}$ <sup>47</sup>. A recent quantum mechanical calculation study of a series of model dipeptides and Ala-X and X-Ala sequences (where X is any amino acid) in both  $\alpha$ -helical and  $\beta$ -sheet conformations showed that the principal values of the tensor were significantly affected by hydrogen bonding at both the carbonyl group and the N-H bond, by the adjacent residues in the polypeptide sequence, and by backbone conformation<sup>48</sup>. Interestingly, the magnitudes of the changes in the orientation of the tensor due to these effects were found to be insignificant compared to the changes in the principal

values. All of these calculations indicate that the variability of  $^{15}\text{N}$  CSA's observed in peptides may not be fully representative of  $^{15}\text{N}$  CSA's in proteins.

Recently, several methods for measurement of  $^{15}\text{N}$  chemical shift anisotropies in solution have been proposed<sup>50-56</sup>. Thus far, residue-specific solution measurements of the  $^{15}\text{N}$  CSA from relaxation measurements have been reported in three proteins: ubiquitin<sup>50-52,54-57</sup>, ribonuclease H<sup>53</sup>, and a small alpha helical protein, C12A-p8<sup>MTCP1</sup><sup>58-60</sup>. These measurements make use of a variety of auto-, cross- and cross-correlation relaxation rates to determine the CSA. In all of these studies, the precision of the available relaxation data has been the limiting factor in determining if the distribution of  $^{15}\text{N}$  CSA's measured in solution for proteins agrees with that which has been reported from solid-state NMR studies on peptides. Specifically in question is whether or not there is significant variability in the  $^{15}\text{N}$  CSA from residue-to-residue (or site-to-site) within proteins, if this variability can be related to protein structure and/or chemistry, and if this relationship matches that predicted by quantum mechanical calculations.

Finding answers to these questions about the  $^{15}\text{N}$  CSAs also has important implications to our understanding of protein dynamics. Though important dynamic contributions to the free energies of reactions of biomolecules need not necessarily be the result of motion on the same timescale as the reaction, most biological interest is clearly focused on motions on the so-called "functional" timescale (i.e. on the  $\mu\text{s}$ -ms timescale range) where most biochemical processes occur. However, most simulation techniques (i.e. molecular dynamics) currently cannot adequately sample these long timescale motions and there is no way to effectively calibrate simulations on these

timescales. The calibration of molecular dynamics simulation of motion in proteins must therefore come from comparison of predicted fast (ps-ns) timescale motions with experiment.  $^{15}\text{N}$  NMR spin relaxation is an excellent method for experimental measurement of motions in proteins on this timescale, however CSA variability makes a significant contribution to experimental dynamic parameters derived from relaxation studies and, due to the computational cost of quantum mechanical calculation of large molecules, the CSA variability cannot be accurately simulated in even small proteins with reasonable computation times. Therefore quantitative calibration of MD simulation with experimental parameters derived from  $^{15}\text{N}$  spin relaxation requires experimental determination of the variability of the  $^{15}\text{N}$  CSA.

Finally, knowledge of the  $^{15}\text{N}$  CSA values is important for the TROSY (Transverse Relaxation Optimized Spectroscopy) technique which reduces  $R_2$  relaxation by making use of the component of a  $^{15}\text{N}$ - $^1\text{H}$  multiplet in which the CSA and dipolar relaxation mechanisms (see Chapter 2) partially cancel each other. Reduced  $R_2$  relaxation results in narrower and higher intensity signals in NMR spectra since the linewidths of NMR signals are proportional to  $R_2$ . The TROSY technique is of particular utility in large molecules where the signals are significantly broadened because  $R_2$  is proportional to the molecule's molecular weight.

### **1.3 Scope of Present Work**

We would like to use  $^{15}\text{N}$  NMR spin relaxation experiments measured at several field strengths to quantitatively evaluate the amplitudes and timescales of motions in a small protein, the third immunoglobulin binding domain of protein G (GB3) on the ps-ms timescale. Since the  $^{15}\text{N}$  CSA is a significant contribution to spin

relaxation, and the range of residue-to-residue variation in this parameter is still in question, it is necessary to measure the site-specific values of the  $^{15}\text{N}$  CSA in GB3 prior to a quantitative analysis of the local motion. Currently solution NMR relaxation rates are the only method for experimental determination of the  $^{15}\text{N}$  CSA at all backbone sites in a protein. Such methods for determination of this parameter measure the field dependence of combinations of autorelaxation and/or CSA/dipolar cross-correlation relaxation rates, or use combinations of many cross-correlation rates, and it is of interest to see if CSAs measured using different methods agree. Since conformational exchange contributions to  $R_2$  (in the fast exchange limit) have the same field dependence as terms proportional to the  $^{15}\text{N}$  CSA, it is necessary to unambiguously identify residues involved in such motions and to quantify the magnitude of the exchange contributions. Furthermore, since the overall tumbling of a molecule is the strongest generator of spin relaxation, and since we would like to use both methods for CSA determination that involve particular models of motion and model-independent methods, any anisotropic motion must be correctly deconvolved from relaxation rates prior to the analysis of the  $^{15}\text{N}$  CSAs and prior to analysis of the local motion. Any neglected anisotropy of the overall tumbling could potentially be misattributed to the CSA or to local motion, by a Lipari-Szabo analysis.

#### **1.4 Specific Aims**

$^{15}\text{N}$  CSA/dipolar cross-correlation rates (CCRs) will be of importance to three aspects of the above general project. These rates are important for identifying residues involved in conformational exchange motions, for determining the overall diffusion tensor of the molecule, and for measurement of site-specific CSAs.

Currently available methods for measurement of  $^{15}\text{N}$  CSA/dipolar cross-correlation rates (CCRs) in proteins rely on two (or more) experiments to measure build-up and decay of the involved coherences. Such “indirect” methods for measuring the CCR are potentially subject to inaccuracy due to experimental differences (in pulses or delays) between the two (or more) experiments. Direct methods, which measure the CCR from the ratio of the two components of the  $^{15}\text{N}$  doublet in a H-coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC-type experiment would not be subject to such inaccuracy. The application of such methods to biological macromolecules, however, could be complicated by signal overlap in the coupled 2D spectra, which may be particularly severe in the case of HSQC spectra of macromolecules with molecular weights greater than 10 kDa. Therefore schemes for spectral simplification that can be applied to H-coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQCs are necessary, and these schemes must be tested within the context of CCR measurement to ensure that such schemes can be safely applied without introducing any scaling factors to measured CCRs. Here we present two such methods and test them on the GB3 domain: the IPAP method of spectral simplification proposed by Ottiger and Bax <sup>61</sup>, and selection of the coherences corresponding to the individual doublet components before the CCR relaxation period by means of the spin-state selective S<sup>3</sup>E element proposed by Sørensen et al. <sup>62</sup>.

The overall rotational diffusion of the molecule is a potential source of anisotropic motion which might have a site-specific contribution to  $^{15}\text{N}$  relaxation rates that could be misattributed to site-specific variation in the  $^{15}\text{N}$  CSA by model-dependent methods for CSA determination or to conformational exchange motion in a standard Lipari-Szabo analysis of local motion. The overall rotational diffusion tensor

and correlation time of the GB3 domain will therefore be determined from  $^{15}\text{N}$  relaxation rates ( $R_1$ ,  $R_2$ ,  $^{15}\text{N}\{^1\text{H}\}$  NOE) and from cross-correlation rates ( $h_z$  and  $h_{xy}$ , see Chapter 4). GB3 represents a particularly stable protein domain, with virtually no tendency to aggregate at high concentration, therefore the measured rotational diffusion parameters can be compared with the predictions of theoretical hydrodynamic models.

A previous analysis of  $^{15}\text{N}$  relaxation rates in the GB1 domain reported elevated  $R_2$  values in the protein's  $\alpha$ -helix and attributed this elevation to exchange broadening. However, this analysis assumed an isotropic model for the overall diffusion of the GB1 domain, and it has since been shown<sup>44,63,64</sup> that neglecting significant rotational anisotropy can result in spurious identification of conformational exchange motion. A precursory examination of the protein data bank structure of the GB3 molecule reveals that the molecule is somewhat elongated in the direction roughly parallel to that of the helical axis of its  $\alpha$ -helix. The principal values of the inertia tensor of GB3 calculated from the structure display anisotropy of  $I_{xx}/I_{zz} = 1.8$ , therefore it is likely that the diffusional anisotropy is significant. Furthermore the principal axis of the highly axially symmetric inertia tensor lies roughly parallel to the helical axis of the  $\alpha$ -helix. To a very rough approximation, it is expected that the symmetry axis of the diffusion tensor be co-linear to the symmetry axis of the inertia tensor; such a condition would result in elevated  $R_2$  rates in the helix, since  $R_2$  is inversely proportional to the rate of tumbling (on the timescale expected for a molecule of GB3's size), and, since the NH vectors in an  $\alpha$ -helix are oriented approximately parallel to the helical axis and thus approximately parallel to the axis

of fast rotational diffusion. The diffusion and inertia tensors of GB3 and GB1 are expected to be very similar (the two domains have 96% sequence homology and 89% sequence identity), therefore it is possible that the elevated  $R_2$  rates observed in the helix of GB1 were due to rotational anisotropy rather than exchange broadening. This will be explored since, as mentioned above, it is necessary to quantify all exchange motion in the GB3 domain prior to an analysis of site-specific values of the  $^{15}\text{N}$  CSA.

A comprehensive study of the  $^{15}\text{N}$  chemical shielding anisotropy in the GB3 domain will be conducted using a combination of  $^{15}\text{N}$  relaxation and  $^{15}\text{N}$  CSA/dipolar cross-correlation measurements over a range of magnetic fields. This study will use both model-of-motion-independent methods and a method which involves a fit to a Lipari-Szabo “model-free” approximation of the spectral density function. The robustness of deriving the CSA from such a fit will be examined. The variability of the  $^{15}\text{N}$  CSA from each of these methods will be determined taking into account their different experimental uncertainties.

Novel  $^{15}\text{N}$  relaxation rates in a protein in  $\text{D}_2\text{O}$  from direct  $^{15}\text{N}$  detection will be presented and demonstrated on the GB3 domain. The proposed direct  $^{15}\text{N}$ -detection experiments offer increased relative sensitivity to  $^{15}\text{N}$  CSA values and could provide a useful tool for accurate measurements of these parameters in proteins. In addition, by sampling the spectral density function at the frequencies  $\omega_D$  and  $\omega_D \pm \omega_N$  these measurements may provide potentially useful information about protein motions in the nanosecond time range not available from the conventional measurements in NH systems. Quantum mechanical *ab initio* calculations at the B3LYP6-311+G(2d,p) level will be conducted on the N-methylacetamide molecule in



the gas phase to explore the sensitivity of the  $^{15}\text{N}$  chemical shielding anisotropy to isotopic substitution of the directly bound hydron.

Finally, motional parameters describing the amplitude and timescales of motions of the peptide planes in the backbone the GB3 domain will be estimated, using the measured diffusion tensor and chemical shift anisotropies. These motional parameters will be compared with motional parameters derived assuming isotropic diffusion and standard values for the  $^{15}\text{N}$  CSA.

## Chapter 2: NMR Relaxation Theory

Phenomenological descriptions of spin-relaxation accompanied the very first demonstrations of nuclear magnetic resonance<sup>24 8</sup>. Spin-relaxation determines the intensities and linewidths of signals in NMR spectra, and many of the inventors of NMR<sup>8,24,65-68</sup>, addressed various issues related to causes of these phenomena. Bloembergen, Purcell and Pound first presented the notion of “motional narrowing”—the idea that the fast tumbling of molecules explains the sharper lines in NMR spectra in gasses and liquids compared to solid-state NMR<sup>8</sup>. Van Vleck further developed this theory<sup>69</sup>. Overhauser explained phenomenologically the dependence of the steady state NOE on motion<sup>66</sup> and Pines and Slichter further developed this theory<sup>68</sup>. Abragam and Pound<sup>65</sup> and Solomon<sup>67</sup> developed expressions for  $R_1$  and  $R_2$ , respectively, for two spin systems, modeling the spin-system as reorienting randomly and isotropically. Woessner extended the expressions for  $R_1$  and  $R_2$  to describe systems in which the molecular tumbling was anisotropic, and introduced the possibility of internal (local) motions of the inter-spin vector with respect to the molecular frame<sup>70 71</sup>. Lipari and Szabo developed a very useful description of spin-relaxation which allows characterization of amplitudes and timescales of motions faster than the overall tumbling at each site in a molecule by two motional-model-independent parameters<sup>27</sup>; as will be discussed further here and in subsequent chapters, these model-free parameters provide a physical picture of molecular motions.

Here I illustrate how  $^{15}\text{N}$  spin relaxation rates in  $^{15}\text{N}$ - $^1\text{H}$  isolated spin pairs can be derived from the so called “master equation” of semi-classical spin relaxation

theory. The treatment here follows closely the work of Slichter<sup>72</sup>, as well as two recent reviews<sup>34,73</sup>. In final form, the expressions for the  $^{15}\text{N}$  relaxation rates (Eqs. 2.5.1-2.5.10) contain information about the motions of  $^{15}\text{N}$ - $^1\text{H}$  bonds through their sampling of the spectral density functions. The  $^{15}\text{N}$  spin relaxation rates therefore, contain information about the dynamics of the polypeptide chain.

### **2.1 The Liouville-von Neumann Equation**

The evolution of a spin-system, considered quantum mechanically, is described by a density operator,  $\sigma$ , with time-dependence given by the Liouville-von Neumann equation:

$$\frac{d}{dt}\sigma = \frac{-i}{\hbar}[\text{H}, \sigma] \quad (2.1.1)$$

where the Hamiltonian can be split into two parts:

$$\text{H} = \text{H}_0 + \text{H}_1(t) \quad (2.1.2)$$

the first of which,  $\text{H}_0$  is time-independent, and the second,  $\text{H}_1(t)$  is a randomly fluctuating perturbation to  $\text{H}_0$ , such that  $\overline{\text{H}_1(t)} = 0$  where the bar indicates a time-average. In the so-called interaction representation, the Liouville-von Neumann equation is:

$$\frac{d}{dt}\tilde{\sigma} = \frac{-i}{\hbar}[\tilde{\text{H}}_1(t), \tilde{\sigma}], \quad (2.1.3)$$

where the bar represents a time average for one member of the ensemble of spins (or according to the ergodic hypothesis equivalently represents an ensemble average over all the spins). Additionally, any operator in the interaction frame,  $\tilde{Q}$  has the relationship to  $Q$  in the stationary frame given by:

$$\tilde{Q} = e^{\frac{i}{\hbar} H_0 t} Q e^{-\frac{i}{\hbar} H_0 t}. \quad (2.1.4)$$

The solution to the Liouville-von Neumann equation is given by:

$$\tilde{\sigma}(t) = \tilde{\sigma}(0) - \frac{i}{\hbar} \int_0^t \overline{[\tilde{H}_1(t'), \tilde{\sigma}(t')] dt'}. \quad (2.1.5)$$

However this cannot be solved in this form as  $\tilde{\sigma}(t')$  in the commutator under the integral is unknown. It can be approximated by:

$$\tilde{\sigma}(t) = \tilde{\sigma}(0) - \frac{i}{\hbar} \int_0^t \overline{[\tilde{H}_1(t'), \tilde{\sigma}(0)] dt'} \quad (2.1.6)$$

if  $H_1(t) \ll H_0$  which implies  $\tilde{\sigma}(t)$  changes slowly with time. A second-order approximation is given by:

$$\tilde{\sigma}(t) = \tilde{\sigma}(0) - \frac{i}{\hbar^2} \int_0^t \int_0^{t'} \overline{[\tilde{H}_1(t'), [\tilde{H}_1(t''), \tilde{\sigma}(0)]] dt'' dt'}. \quad (2.1.7)$$

This amounts to a perturbation expansion approach to the solution, placing the entire expression for  $\tilde{\sigma}(t)$  in Eq. 2.1.5, inside the commutator in Eq. 2.1.5.

## **2.2 Hamiltonian for Spin Relaxation**

For  $^{15}\text{N}$ - $^1\text{H}$  spin systems, the time dependent perturbation  $H_1(t)$  to the stationary Hamiltonian experienced by the  $^{15}\text{N}$  nucleus is caused by fluctuation of the dipolar interaction with the  $^1\text{H}$  spin and by fluctuation of its chemical shielding.

$H_1(t)$  is therefore given by two terms, one arising from the dipole-dipole interaction between the  $^{15}\text{N}$  and  $^1\text{H}$  nuclei,  $H_1^{DD}(t)$ , and a second arising from the anisotropy of the nitrogen chemical shielding,  $H_1^{CSA}(t)$ :

$$H_1(t) = H_1^{DD}(t) + H_1^{CSA}(t), \quad (2.2.1)$$

where the dipole-dipole and  $^{15}\text{N}$  chemical shielding anisotropy terms in the Hamiltonian in the laboratory frame (the frame with z-axis in the direction of the external magnetic field,  $B_0$ ) are given by:

$$H_1^{DD}(t) = 2d \left( \vec{H} \cdot \vec{N} - 3(\vec{H} \cdot \hat{r}_{NH})(\vec{N} \cdot \hat{r}_{NH}) \right), \quad (2.2.2)$$

and

$$H^{CSA} = \gamma_N \vec{B}_0 \cdot \underline{\underline{\sigma}} \cdot \vec{N}. \quad (2.2.3)$$

Here  $d$  is the dipole-dipole interaction constant,  $d = -\mu_0 \gamma_H \gamma_N \hbar / (8\pi r_{NH}^3)$  and depends on the gyromagnetic ratios of both  $^{15}\text{N}$  and  $^1\text{H}$ ,  $\gamma_N$  and  $\gamma_H$ , and the distance between the nuclei,  $r_{NH}$ .  $\vec{H}$  and  $\vec{N}$  are the angular momentum operators for the nitrogen and proton spins and  $\hat{r}_{NH}$  is the unit vector along the line connecting the nitrogen and proton nuclei, i.e.  $\hat{r}_{NH} = \frac{\vec{r}_{NH}}{r_{NH}}$ .

In the dipolar interaction frame, the principal axes frame (PAF) with Z axis along the N-H bond,  $H_1^{DD}(t)$  becomes:

$$H_1^{DD}(t) = 2d \left( \vec{H} \cdot \vec{N} - 3H_z N_z \right). \quad (2.2.4)$$

in Eq. 2.2.3  $\underline{\underline{\sigma}}$  is the  $^{15}\text{N}$  chemical shielding tensor which describes the electronic environment of the nitrogen nucleus. If  $\underline{\underline{\sigma}}$  is symmetric it has real eigenvalues and orthogonal eigenvectors and thus the transformation that diagonalizes  $\underline{\underline{\sigma}}$  is a rotation to the chemical shielding interaction frame, the PAF of the chemical shielding tensor, where  $\underline{\underline{\sigma}}$  is diagonal. In this frame,  $\Delta\sigma$  is the anisotropy of this tensor, abbreviated CSA, and given by:<sup>74</sup>

$$|\Delta\sigma| = \left| \sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 - (\sigma_{xx}\sigma_{yy} + \sigma_{xx}\sigma_{zz} + \sigma_{yy}\sigma_{zz}) \right|^{1/2}, \quad (2.2.5)$$

and the rhombicity of the chemical shielding tensor is given by:

$$\xi = \left( \frac{\sigma_{yy} - \sigma_{xx}}{\Delta\sigma} \right), \quad (2.2.6)$$

where  $\sigma_{xx}$ ,  $\sigma_{yy}$ , and  $\sigma_{zz}$  are the eigenvalues of the tensor in its PAF.  $H^{CSA}$  contains both time-independent and time-dependent terms, but only the terms from the anisotropy contribute to relaxation since a nucleus with isotropic chemical shielding does not experience fluctuations in its magnetic environment due to rotation, though fluctuations in the diagonal terms contribute to  $R_{ex}$ .  $H^{CSA}$  can be separated:

$$\begin{aligned} H_{PAF}^{CSA} &= \gamma_N (\sigma_{xx} B_x N_x + \sigma_{yy} B_y N_y + \sigma_{zz} B_z N_z) \\ &= \gamma_N \left( \frac{\sigma_{xx} + \sigma_{yy} + \sigma_{zz}}{3} \vec{B} \cdot \vec{N} + \frac{\Delta\sigma}{3} (2B_z N_z - B_x N_x - B_y N_y) \right). \end{aligned} \quad (2.2.7)$$

Note that the part contributing to relaxation can be written in a form identical to that of the dipolar Hamiltonian (Eq.2.2.4):

$$H_1^{CSA}(t) = -\gamma_N \frac{\Delta\sigma}{3} (\vec{B} \cdot \vec{N} - 3B_z N_z). \quad (2.2.8)$$

The CSA interaction constant,  $c$  is given by:

$$c = \gamma_N B_0 \Delta\sigma / 3 = -\omega_N \cdot \Delta\sigma / 3. \quad (2.2.9)$$

Both of the time-dependent Hamiltonians in Eqs. 2.2.4 and 2.2.7 can be separated as:

$$H_1(t) = \sum_m T_m(N, H) F_m(\theta(t), \varphi(t)) \quad (2.2.10)$$

where the  $T_m$  are spin operators acting on the spin variables (operators  $N$ ,  $H$ ), and the  $F_m$  are spatial functions of the time-dependent angles  $(\theta(t), \varphi(t))$  defining the

orientation of the interaction vector (either the bond vector or the symmetry axis of the CSA tensor) with respect to the laboratory frame (defined as the frame with z-axis along the external field). For both the dipole-dipole and CSA interactions, the functions  $F_m$  are proportional to second rank spherical harmonics  $Y_{2,m}$ , the  $T_m$  and  $F_m$  are given by in Table 2.2.1 below:

| Dipole-dipole interaction:  | CSA interaction:   |
|---|--|
| $H_1^{DD}(t) = \sum_m T_m^{DD} F_m^{DD}(t)$   | $H_1^{CSA}(t) = \sum_m T_m^{CSA} F_m^{CSA}(t)$   |
| $T_m^{DD} = \frac{\mu_o \gamma_H \gamma_N \hbar}{4\pi \langle r^3 \rangle} T_{2,m}^{DD}$ $T_{2,0}^{DD} = \frac{1}{\sqrt{6}} (4I_z S_z - (I_+ S_- + I_- S_+))$ $T_{2,\pm 1}^{DD} = \pm (I_z S_\pm + I_\pm S_z)$ $T_{2,\pm 2}^{DD} = -I_\pm S_\pm$ $F_m^{DD}(t) = \sqrt{4\pi} Y_{2,m}^*(\theta_{DD}(t), \varphi_{DD}(t))$ | $T_m^{CSA} = \frac{\gamma_S \Delta \sigma}{2\sqrt{3}} T_{2,m}^{CSA}$ $T_{2,0}^{CSA} = -\frac{1}{\sqrt{6}} (4B_z S_z - (B_+ S_- + B_- S_+)) = \frac{-4}{\sqrt{6}} B_o S_z$ $T_{2,\pm 1}^{CSA} = \pm (B_z S_\pm + S_z S_\pm) = \pm B_o S_\pm$ $T_{2,\pm 2}^{CSA} = -B_\pm S_\pm = 0$ $F_m^{CSA}(t) = \sqrt{4\pi} Y_{2,m}^*(\theta_{CSA}(t), \varphi_{CSA}(t))$ |

where  $\theta_{DD}(t)$  and  $\varphi_{DD}(t)$  define the orientation of the bond vector (dipole-dipole interaction vector) in the laboratory frame. For an axially symmetric chemical shift anisotropy interaction,  $\theta_{CSA}(t)$  and  $\varphi_{CSA}(t)$  define the orientation of the symmetry axis

of the CSA tensor in the laboratory frame. For a rhombic CSA tensor, the tensor is expressed as the sum of two axially symmetric tensors<sup>75</sup>.

### **2.3 Separation of Spin Operators and Spatial Functions**

In general, any perturbing Hamiltonian can be written as a sum of terms from Table 2.1.1, where the sum over  $i$  indicates over all possible interaction (i.e. dipolar and CSA) terms:

$$H_1(t) = \sum_m \sum_i T_m^i F_m^i(t). \quad (2.3.1)$$

In the interaction representation, each of the  $T_m^i$  spin operators rotates with a characteristic frequency:

$$\langle \alpha | \tilde{T}_m^i | \beta \rangle = \langle \alpha | e^{\frac{i}{\hbar} H_0 t} T_m^i e^{-\frac{i}{\hbar} H_0 t} | \beta \rangle = e^{i(\omega_\alpha - \omega_\beta)t} \langle \alpha | T_m^i | \beta \rangle \quad (2.3.2)$$

so that:

$$\tilde{T}_m^i = e^{i\omega_m^i t} T_m^i. \quad (2.3.3)$$

The random spatial functions,  $F_m^i(t)$  average to zero so that  $\overline{H_1(t)} = 0$ , and have time independent ensemble averages. One can write products of the  $F_m^i(t)$  in terms of time correlation functions:

$$\overline{F_m^i(t) F_{m'}^i(t')} = C_{mm'}(|t - t'|), \quad (2.3.4)$$

where  $C_{mm'}(t)$  has the properties that,  $C_{mm'}(\infty) = 0$  and  $C_{mm'}(\tau) = -C_{mm'}(\tau)$ . In terms of the time correlation functions, Eq. 2.1.7 can be written:

$$\tilde{\sigma}(t) - \tilde{\sigma}(0) = -\frac{i}{\hbar^2} \sum_{m,m'} \sum_i [T_m^i, [T_{m'}^{j*}(\tilde{\sigma}(0) - \tilde{\sigma}_{eq})]] \times \int_0^t \int_0^{t'} C_{mm'}(|t' - t''|) e^{i(\omega_m^i t' - \omega_{m'}^j t'')} dt'' dt', \quad (2.3.5)$$



where the integral in the instance  $m=m'$  is performed by a change of variables,

$$\tau = t' - t'':$$

$$\int_0^t \int_0^{t'} C_{mm}(|t' - t''|) e^{i\omega_m^i(t' - t'')} dt'' dt' = \int_0^t \int_0^\tau C_{mm}(\tau) e^{i\omega_m^i \tau} dt' d\tau = \int_0^t (t - \tau) C_{mm}(\tau) e^{i\omega_m^i \tau} d\tau. \quad (2.3.6)$$

If  $C_{mm}$  decays with some characteristic time,  $\tau_c$ , so that  $C_{mm}(\tau < \tau_c) \approx C_{mm}(0)$ , and

$C_{mm}(\tau \gg \tau_c) \approx 0$ , and if the time,  $t$ , can be chosen so that the change in  $\tilde{\sigma}$  during  $t$  is

small, and that  $\tau_c < t$  then:

$$\int_0^t (t - \tau) C_{mm}(\tau) e^{i\omega_m^i \tau} d\tau \approx t \int_0^\infty C_{mm}(\tau) e^{i\omega_m^i \tau} d\tau = \frac{t}{2} \int_{-\infty}^{+\infty} C_{mm}(\tau) e^{i\omega_m^i \tau} d\tau = \frac{t}{2} J_{mm}(\omega_m^i) \quad (2.3.7)$$

This then serves as a definition of the spectral density function,  $J_{mm}^i(\omega_m^i)$ :

$$J_{mm}^i(\omega_m^i) = \int_{-\infty}^{+\infty} C_{mm}(\tau) e^{i\omega_m^i \tau} d\tau = 2 \int_0^\infty C_{mm}(\tau) e^{i\omega_m^i \tau} d\tau. \quad (2.3.8)$$

For  $m \neq m'$ ,

$$\begin{aligned} \int_0^t \int_0^{t'} C_{mm'}(|t' - t''|) e^{i(\omega_m^i t' - \omega_{m'}^i t'')} dt'' dt' &= \int_0^t \int_\tau^t C_{mm'}(\tau) e^{i\omega_m^i t'} e^{-i\omega_{m'}^i (t' - \tau)} dt' d\tau = \int_0^t \int_\tau^t C_{mm'}(\tau) e^{i(\omega_m^i - \omega_{m'}^i) t'} e^{i\omega_{m'}^i \tau} dt' d\tau \\ &= \frac{1}{i(\omega_m - \omega_{m'})} \int_0^t \left( e^{i(\omega_m - \omega_{m'}) t} - e^{i(\omega_m - \omega_{m'}) \tau} \right) C_{mm'}(\tau) e^{i\omega_{m'}^i \tau} d\tau = \frac{1}{2i(\omega_m - \omega_{m'})} \left( e^{i(\omega_m - \omega_{m'}) t} J_{mm'}(\omega_{m'}) - J_{mm'}(\omega_m) \right) \end{aligned} \quad (2.3.9)$$

If the frequency difference between the two states  $m$  and  $m'$  is large, the terms proportional to  $J_{mm}$  will be larger than terms proportional to  $J_{mm'}$ . In terms of these spectral density functions, Eq. 2.3.5 now gives the master equation for a density

operator subject to a randomly fluctuating perturbation:

$$\frac{d\tilde{\sigma}}{dt} \approx \frac{\tilde{\sigma}(t) - \tilde{\sigma}(0)}{t} = -\frac{i}{\hbar^2} \frac{1}{2} \sum_m [T_m^i, [T_m^{j*} (\tilde{\sigma}(0) - \tilde{\sigma}_{eq})]] J_{mm}(\omega_m), \quad (2.3.10)$$

where we have chosen the evolution time,  $t$ , such that the change in the density matrix is small. Then the evolution of any observable,  $Q$ , is given by:

$$\begin{aligned} \frac{d\langle Q \rangle}{dt} &= \frac{d}{dt} \text{Tr}(Q\tilde{\sigma}) = \text{Tr}(Q \frac{d\tilde{\sigma}}{dt}) = -\frac{1}{\hbar^2} \frac{1}{2} \sum_m \text{Tr}(Q [T_m^i [T_m^{j*} (\tilde{\sigma}(0) - \tilde{\sigma}_{eq})]]) J_{mm}(\omega_m) \\ &= -\frac{1}{\hbar^2} \frac{1}{2} \sum_m \text{Tr}([Q, T_m^i] [T_m^{j*} (\tilde{\sigma}(0) - \tilde{\sigma}_{eq})]) J_{mm}(\omega_m) \\ &= -\frac{1}{\hbar^2} \frac{1}{2} \sum_m \left( \sum_{i,j} \langle [[Q, T_m^i], T_m^{j*}] \rangle - \langle [[Q, T_m^i], T_m^{j*}] \rangle_{eq} \right) J_{mm}(\omega_m) \end{aligned} \quad (2.3.11)$$

The  $i=j$  terms are autocorrelation functions and give rise to autorelaxation rates while the  $i \neq j$  terms give rise to interference effects which are termed cross-correlations between relaxation pathways and contribute to cross-correlation rates (CCRs).

## **2.4 Correlation and Spectral Density Functions**

For the general case of mono-exponential decay of the correlation function with one time-constant, as:

$$C_{mm}(\tau) = \overline{F_m(t) F_m^*(t + \tau)} = \overline{|F_m|^2} e^{-\tau/\tau_c} \quad (2.4.1)$$

the spectral density function can be calculated as:

$$J_m(\omega) = 2 \int_{-\infty}^{\infty} C_{mm}(\tau) e^{i\omega\tau} d\tau = 2 \overline{|F_m|^2} \int_{-\infty}^{\infty} e^{-\frac{\tau}{\tau_c}} e^{i\omega\tau} d\tau = 2 \overline{|F_m|^2} \frac{1}{\frac{1}{\tau_c} - i\omega} = 2 \overline{|F_m|^2} \left( \frac{\tau_c}{1 + \omega^2 \tau_c^2} + i \frac{\omega \tau_c}{1 + \omega^2 \tau_c^2} \right) \quad (2.4.2)$$

Therefore, when the decay of the correlation function is described by a mono-exponential function with characteristic time,  $\tau_c$ , the spectral density function is given by:

$$J_m(\omega) = 2 \overline{|F_m|^2} \frac{\tau_c}{1 + \omega^2 \tau_c^2} \quad (2.4.3)$$

This then, is the spectral density for the reorientation of an interaction with respect to the magnetic field. Eq. 2.4.3 can be used to describe relaxation in a  $^{15}\text{N}$ - $^1\text{H}$  spin pair, where the spins are rigidly attached to a spherical molecule, with no motion of the relaxation interaction vector (the NH bond or the symmetry axis of the CSA interaction) except for that due to the isotropic overall rotation. In this case, the bond vector is randomly distributed in space and  $\overline{|F_m|^2} = \frac{1}{5}$  as  $\overline{F_m F_{m'}^*} = \delta_{mm'} \frac{1}{5}$  due to the orthogonality properties of the spherical harmonics. In general, the spherical harmonic functions (as in table 2.2.1) which appear in the spatial functions are related to the Wigner rotation matrices by the relationship:

$$Y_l^{m'}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} D_{m'0}^{(l)}(\Omega_R), \quad (2.4.4)$$

where  $D_{m,n}^{(l)}(\Omega_R)$  is a Wigner rotation matrix<sup>76</sup> which specify transformations of spherical tensors under rotations, with  $\Omega_R$  representing the set of Euler angles<sup>77</sup> that describe the rotation ( $\Omega_R$  denotes the Euler angles  $\theta, \phi, \psi$ ).

$$C(\tau) = \left\langle D_{q,0}^{(2)*}(\Omega_{lab \rightarrow inst}^t) D_{q,0}^{(2)}(\Omega_{lab \rightarrow inst}^{t+\tau}) \right\rangle_{lab \rightarrow inst} \quad (2.4.5)$$

or resetting the zero of time:

$$C(\tau) = \left\langle D_{q,0}^{(2)*}(\Omega_{lab \rightarrow inst}^0) D_{q,0}^{(2)}(\Omega_{lab \rightarrow inst}^\tau) \right\rangle_{lab \rightarrow inst} \quad (2.4.6)$$

where the  $\Omega_{lab \rightarrow inst}$  are the set of three Euler angles that specify the orientation of the unit vector describing the interaction (i.e. the symmetry axis of the CSA tensor or the NH bond vector for the dipolar interaction) in the laboratory frame. Here the  $\Omega_{lab \rightarrow inst}$  can specify either the transformation of the coordinates of the bond vector from the lab frame to its instantaneous orientation (for dipolar interactions), or the transformation of the symmetry axis of a CSA tensor from the lab frame to its instantaneous orientation (for the CSA interaction). The brackets denote an ensemble average over all initial,  $\Omega_{lab \rightarrow inst}^0$ , and final,  $\Omega_{lab \rightarrow inst}^t$ , orientations.

For conceptual convenience, the Woessner<sup>70 71</sup> and Lipari-Szabo<sup>27</sup> descriptions involve factorization of the correlation function into terms representing different transformations, e.g.:

$$C(\tau) = \sum_{m=-2}^2 \sum_{n=-2}^2 \left\langle D_{q,m}^{(2)*}(\Omega_{lab \rightarrow loc}^0) D_{q,n}^{(2)}(\Omega_{lab \rightarrow loc}^t) D_{m,0}^{(2)*}(\Omega_{loc \rightarrow inst}^0) D_{n,0}^{(2)}(\Omega_{loc \rightarrow inst}^t) \right\rangle, \quad (2.4.7)$$

where the transformation from the lab frame to the instantaneous frame has been conceptually broken into the transformation from the lab frame to the time-averaged molecular frame (designated “*loc*”, assumed to be represented by the protein data bank structure of the molecule) and another transformation from the molecular frame to the instantaneous orientation at time  $t$ , designated “*inst*”.

At this point an important approximation is made, namely that the overall tumbling and local motion of the interaction vectors are essentially independent. For isotropic overall diffusion and local motion, this amounts to re-writing Eq. 2.4.7 as:

$$C(\tau) = C_o(\tau) C_l(\tau) = \sum_{m=-2}^2 \sum_{n=-2}^2 \left\langle D_{q,m}^{(2)*}(\Omega_{lab \rightarrow loc}^0) D_{q,n}^{(2)}(\Omega_{lab \rightarrow loc}^t) \right\rangle \left\langle D_{m,0}^{(2)*}(\Omega_{loc \rightarrow inst}^0) D_{n,0}^{(2)}(\Omega_{loc \rightarrow inst}^t) \right\rangle \quad (2.4.8)$$

where:

$$C_o(\tau) = \left\langle D_{q,m}^{(2)*}(\Omega_{lab \rightarrow loc}^0) D_{q,n}^{(2)}(\Omega_{lab \rightarrow loc}^t) \right\rangle = \frac{1}{5} e^{-\tau/\tau_c} \delta_{mn}, \quad (2.4.9)$$

with the correlation for internal motion given by:

$$C_I(\tau) = \left\langle D_{m,0}^{(2)*}(\Omega_{loc \rightarrow inst}^0) D_{n,0}^{(2)}(\Omega_{loc \rightarrow inst}^t) \right\rangle \quad (2.4.10)$$

In terms of the more familiar spherical harmonics, and in the instance of isotropic overall diffusion, Eq. 2.4.8 can now be simplified to:

$$C(\tau) = \frac{4\pi}{5} \left\langle Y_{2,0}^*(\theta(0), \varphi(0)) Y_{2,0}(\theta(\tau), \varphi(\tau)) \right\rangle. \quad (2.4.11)$$

In the case of anisotropic overall rotation (where the effective correlation time about one direction in space is faster than others) the correlation function does not decay as a mono-exponential function with one correlation time. For anisotropic overall diffusion and local motion, the correlation function can be factored as:

$$C(\tau) = \sum_{m=-2}^2 \sum_{n=-2}^2 \sum_{k=-2}^2 \sum_{l=-2}^2 \left\langle D_{q,m}^{(2)*}(\Omega_{lab \rightarrow dif}^0) D_{q,n}^{(2)}(\Omega_{lab \rightarrow dif}^t) \right\rangle \\ \times \left\langle D_{m,k}^{(2)*}(\Omega_{dif \rightarrow loc}^0) D_{n,l}^{(2)}(\Omega_{dif \rightarrow loc}^t) \right\rangle \left\langle D_{k,0}^{(2)*}(\Omega_{loc \rightarrow inst}^0) D_{l,0}^{(2)}(\Omega_{loc \rightarrow inst}^t) \right\rangle. \quad (2.4.12)$$

Here, “dif” designates the PAF of the overall rotational diffusion tensor of the molecule. When the local motion is ignored, the coordinates in the average molecular frame  $(x, y, z)$  are related to the coordinates in the diffusion tensor frame  $(x_d, y_d, z_d)$  by the simple rotation:

$$\begin{bmatrix} x_d \\ y_d \\ z_d \end{bmatrix} = R(\Phi, \Theta, \Psi) \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (2.4.13)$$

where the rotation matrix,  $R(\Phi, \Theta, \Psi)$ , is given by:

$$R(\Phi, \Theta, \Psi) = \begin{bmatrix} \cos\Phi \cos\Theta \cos\Psi - \sin\Phi \sin\Psi & \sin\Phi \cos\Theta \cos\Psi + \cos\Phi \sin\Psi & -\sin\Theta \sin\Psi \\ -\cos\Phi \cos\Theta \cos\Psi - \sin\Phi \sin\Psi & -\sin\Phi \cos\Theta \cos\Psi + \cos\Phi \sin\Psi & \sin\Theta \sin\Psi \\ \cos\Phi \sin\Theta & \sin\Phi \sin\Theta & \cos\Theta \end{bmatrix} \quad (2.4.14)$$

where  $\Omega_{loc \rightarrow dif} = \{\Phi, \Theta, \text{ and } \Psi\}$  are the Euler angles relating the PAF of the diffusion tensor to the PAF of the averaged molecular reference frame. In the case of axial symmetry of the overall diffusion, Woessner showed that the spectral density function is given by:

$$J(\omega) = \frac{2}{5} \sum_{i=1}^3 \left[ \frac{A_i^{ax} D_i^{ax}}{(D_i^{ax})^2 + \omega^2} \right] \quad (2.4.15)$$

Where the  $D_i^{ax}$  and  $A_i^{ax}$  are given by:

$$\begin{aligned} D_1^{ax} &= 5D_{\perp} + D_{\parallel} \\ D_2^{ax} &= 2D_{\perp} + 4D_{\parallel} , \\ D_3^{ax} &= 6D_{\perp} \end{aligned} \quad (2.4.16)$$

and

$$\begin{aligned} A_1^{ax} &= 3z_d^2(1 - z_d^2) \\ A_2^{ax} &= \frac{3}{4}(1 - z_d^2)^2 \\ A_3^{ax} &= \frac{1}{4}(3z_d^2 - 1)^2 \end{aligned} \quad (2.4.17)$$

The  $D_{\parallel}$  and  $D_{\perp}$  are the principal components of the rotational diffusion tensor.

For a molecule which may be modeled by a completely anisotropic diffusion tensor, the spectral density is given by:

$$J(\omega) = \frac{2}{5} \sum_{i=1}^5 A_i \left[ \frac{A_i D_i}{D_i^2 + \omega^2} \right], \quad (2.4.18)$$

where

$$\begin{aligned} D_1 &= 4D_{xx} + D_{yy} + D_{zz} \\ D_2 &= D_{xx} + 4D_{yy} + D_{zz} \\ D_3 &= D_{xx} + D_{yy} + 4D_{zz} \\ D_4 &= 6D_{iso} + 6\sqrt{D_{iso}^2 - D^2} \\ D_5 &= 6D_{iso} - 6\sqrt{D_{iso}^2 - D^2} \end{aligned} \quad (2.4.19)$$

with  $D_{iso} = (D_{xx} + D_{yy} + D_{zz})/3$  and  $D^2 = (D_{xx}D_{yy} + D_{yy}D_{zz} + D_{xx}D_{zz})/3$  where by

definition ( $D_{xx} \leq D_{yy} \leq D_{zz}$ ). The  $A_i$ 's in Equation 2.4.18 are functions of the angles

between the NH vectors and the principal axes of the diffusion tensor, such that:

$$\begin{aligned} A_1 &= 3y_d^2 z_d^2 \\ A_2 &= 3x_d^2 z_d^2 \\ A_3 &= 3x_d^2 y_d^2 \\ A_4 &= \frac{1}{4}[3(x_d^4 + y_d^4 + z_d^4) - 1] - \frac{1}{12}[\delta_x(3x_d^4 + 6y_d^2 z_d^2 - 1) + \delta_y(3y_d^4 + 6x_d^2 z_d^2 - 1) + \delta_z(3z_d^4 + 6y_d^2 x_d^2 - 1)] \\ A_5 &= \frac{1}{4}[3(x_d^4 + y_d^4 + z_d^4) - 1] + \frac{1}{12}[\delta_x(3x_d^4 + 6y_d^2 z_d^2 - 1) + \delta_y(3y_d^4 + 6x_d^2 z_d^2 - 1) + \delta_z(3z_d^4 + 6y_d^2 x_d^2 - 1)] \end{aligned} \quad (2.4.20)$$

$$\text{where } \delta_j = (D_{jj} - D_{iso})/\sqrt{D_{iso}^2 - D^2} \quad (j = x, y, z). \quad (2.4.21)$$

In the Lipari-Szabo approximation (also termed the “model free”

approximation) the correlation time for internal motion,  $C_I(\tau)$ , is approximated with

a single exponential decay which agrees with Eq. 2.4.10 at  $\tau = 0$  and at  $\tau = \infty$ , and

bounds a region of the same area as bounded by Eq. 2.4.10 such that:

$$C_I(\tau) = \frac{1}{5} e^{-\tau/\tau_c} \left( C_I(\tau \rightarrow \infty) + (C_I(0) - C_I(\tau \rightarrow \infty)) e^{-\tau/\tau_c} \right). \quad (2.4.22)$$

The  $\tau=0$  limit of Eq. 2.4.10 is:

$$C_I(0) = 1 \quad (2.4.23)$$

The  $\tau \rightarrow \infty$  limit depends on the specific model of motion, but is assigned such that  $C_I(\infty) \equiv S^2$ , so that, in the Lipari-Szabo expansion with isotropic overall motion,  $C(\tau)$  is approximated as:

$$C_I(\tau) = \frac{1}{5} e^{-\tau/\tau_c} \left( S^2 + (P_2(\cos \theta) - S^2) e^{-\tau/\tau_c} \right) \quad (2.4.22)$$

Here  $S$  is called the order parameter, and describes the degree of spatial restriction of the interaction vector such that  $S=1$  corresponds to a single orientation and  $S=0$  corresponds to sampling of  $180^\circ$  of angular space, while  $\tau_c$  is the local correlation time for the motion.

## **2.5 Relaxation Rates**

If decoupling is properly applied during the relaxation delay and all interference terms are suppressed, relaxation rates,  $R_1$  and  $R_2$  contain only terms due to autocorrelation functions of the dipolar and CSA relaxation mechanisms. This means that the dipolar and CSA time-dependent perturbations to the Hamiltonian can be considered separately, and that the rates can be expressed as sums of spectral density functions for these interactions ( $J^{DD}(\omega)$  and  $J^{CSA}(\omega)$ , respectively) multiplied by squared interaction constants ( $d^2$  and  $c^2$ ) with no interference terms. These rates are calculated using equation 2.3.12 where  $Q$  is the  $^{15}\text{N}$  angular momentum operator,  $N_z$  for longitudinal relaxation,  $N_x$  and  $N_y$  for transverse relaxation, and the  $T_m^i$  are



given by Table 2.3.1. The spectral density functions are calculated from the time-dependent angular functions  $F_m^i$  in Table 2.1.1.

For  $^{15}\text{N}$  relaxation rates in systems that can be approximated as two-spin,  $^{15}\text{N}$ - $\{^1\text{H}\}$  spin systems:

$$R_1 = d^2 [J^{DD}(\omega_H - \omega_N) + 6J^{DD}(\omega_H + \omega_N) + 3J^{DD}(\omega_N)] + c^2 3J^{CSA}(\omega_N) \quad (2.5.1)$$

$$R_2 = \frac{d^2}{2} [J^{DD}(\omega_H - \omega_N) + 6J^{DD}(\omega_H) + 6J^{DD}(\omega_H + \omega_N) + 4J^{DD}(0) + 3J^{DD}(\omega_N)] + c^2 [2J^{CSA}(0) + \frac{3}{2}J^{CSA}(\omega_N)] + R_{ex} \quad (2.5.2)$$

When the spectral density function for the dipolar interaction,  $J^{DD}(\omega)$ , and the spectral density function for the CSA interaction,  $J^{CSA}(\omega)$ , are similar enough to be approximated by one spectral density function,  $J(\omega)$ , these equations reduce to their more familiar forms:

$$R_1 = d^2 [J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)] + (d^2 + c^2) 3J(\omega_N) \quad (2.5.3)$$

$$R_2 = \frac{d^2}{2} [J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + (d^2 + c^2) [4J(0) + 3J(\omega_N)] + R_{ex} \quad (2.5.4)$$

For a fully anisotropic CSA tensor, these expressions are:

$$R_1 = d^2 [J^{DD}(\omega_H - \omega_N) + 6J^{DD}(\omega_H + \omega_N) + 3J^{DD}(\omega_N)] + \quad (2.5.5)$$

$$\frac{\omega_N^2}{3} [(\sigma_{zz} - \sigma_{xx})^2 J_{zz}^{CSA}(\omega_N) + (\sigma_{yy} - \sigma_{xx})^2 J_{yy}^{CSA}(\omega_N) - (\sigma_{zz} - \sigma_{xx})(\sigma_{yy} - \sigma_{xx}) J_{zy}^{CSA}(\omega_N)]$$

$$R_2 = \frac{d^2}{2} [J^{DD}(\omega_H - \omega_N) + 6J^{DD}(\omega_H) + 6J^{DD}(\omega_H + \omega_N) + 4J^{DD}(0) + 3J^{DD}(\omega_N)] +$$

$$\begin{aligned}
& \frac{2\omega_N^2}{9}[(\sigma_{zz} - \sigma_{xx})^2 J_{zz}^{CSA}(0) + (\sigma_{yy} - \sigma_{xx})^2 J_{yy}^{CSA}(0) - (\sigma_{zz} - \sigma_{xx})(\sigma_{yy} - \sigma_{xx})J_{zy}^{CSA}(0)] + \\
& \frac{\omega_N^2}{6}[(\sigma_{zz} - \sigma_{xx})^2 J_{zz}^{CSA}(\omega_N) + (\sigma_{yy} - \sigma_{xx})^2 J_{yy}^{CSA}(\omega_N) - (\sigma_{zz} - \sigma_{xx})(\sigma_{yy} - \sigma_{xx})J_{zy}^{CSA}(\omega_N)] + R_{ex}
\end{aligned}
\tag{2.5.6}$$

which also reduce to Eqs. 2.5.3-4 when

$$J^{DD}(\omega) \approx J_{xx}^{CSA}(\omega) \approx J_{yy}^{CSA}(\omega) \approx J_{zz}^{CSA}(\omega) \approx J_{zy}^{CSA}(\omega).$$

The longitudinal and transverse  $^{15}\text{N}$  dipole-dipole/CSA cross-correlation rates (henceforth CCRs) will be discussed in Chapter 4. For completeness, these rates are given here in terms of the spectral density functions. The longitudinal,  $\eta_z$ , and transverse,  $\eta_{xy}$ , CCRs are given by:

$$\eta_z = 6dcJ^{DD,CSA}(\omega_N)P_2(\cos\beta), \text{ and} \tag{2.5.7}$$

$$\eta_{xy} = dc[4J^{DD,CSA}(0) + 3J^{DD,CSA}(\omega_N)]P_2(\cos\beta), \tag{2.5.8}$$

for an axially symmetric  $^{15}\text{N}$  CSA tensor, and by:

$$\eta_z = 2\omega_N dJ(\omega_N)[(\sigma_{zz} - \sigma_{yy})P_2(\cos\beta_z) + (\sigma_{yy} - \sigma_{xx})P_2(\cos\beta_x)] \tag{2.5.9}$$

$$\begin{aligned}
\eta_{xy} = & \frac{\omega_N}{3}d[(\sigma_{zz} - \sigma_{xx})P_2(\cos\beta_z)][4J_{zz}^{DD,CSA}(0) + 3J_{zz}^{DD,CSA}(\omega_N)] + \\
& [(\sigma_{yy} - \sigma_{xx})P_2(\cos\beta_x)][4J_{yy}^{DD,CSA}(0) + 3J_{yy}^{DD,CSA}(\omega_N)],
\end{aligned}
\tag{2.5.10}$$

for a fully anisotropic CSA tensor.  $P_2(x)$  is the Legendre polynomial, and  $\beta$  is the angle between the unique axis of an axially symmetric CSA tensor and the NH bond vector, while  $\beta_z$ ,  $\beta_x$  are the intervening angles between the principal axes (z and x) of

the fully anisotropic  $^{15}\text{N}$  CSA tensor and the NH bond vector.  $J^{DD,CSA}(\omega)$  is the cross-correlation spectral density function.

Taken together, the relationships for  $R_1$  and  $R_2$  given by Eqs. 2.5.3-2.5.4 contain terms proportional to  $J(0)$ ,  $J(\omega_N)$ ,  $J(\omega_H)$ ,  $J(\omega_N + \omega_H)$ , and  $J(\omega_N - \omega_H)$ . Due to the inverse dependence on  $\omega^2$  of the spectral density function (cf. Eq. 2.4.3 and 2.4.18), the spectral density functions at high frequencies are small relative to the spectral density functions at low frequencies. It has been shown<sup>31</sup> that since  $\omega_H \approx 10 \cdot \omega_N$ , to a good approximation:

$$J(\omega_H) \approx J(\omega_H + \omega_N) \approx J(\omega_H - \omega_N) \quad (2.5.11)$$

This approximation can be improved by making the assumption

that  $J(\omega) = \lambda_1 / \omega^2 + \lambda_2$  where the first and second terms represent contributions to  $J(\omega)$  from the overall and internal dynamics respectively, and requiring that the relationship between the spectral density at a single “equivalent” frequency,  $\omega_q$ , and  $J(\omega_H \pm \omega_N)$  is given by:

$$5J(\omega_q) = 6J(\omega_H + \omega_N) - J(\omega_H - \omega_N) \quad (2.5.12)$$

This relation can then be recast so that:

$$\omega_q = \{5 / (6 [1 + (\gamma_N / \gamma_H)]^2 - 1 / [1 - (\gamma_N / \gamma_H)]^2)\}^{1/2} \omega_H, \quad (2.5.13)$$

yielding  $\omega_q = 0.870 \omega_H$ . In an analogous manner, it can be shown that<sup>78</sup>:

$$7J(0.921\omega_H) = J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N), \quad (2.5.14)$$

and

$$13J(0.955\omega_H) = J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N). \quad (2.5.15)$$

These results are termed the reduced spectral density approximation<sup>78,79</sup>, under which it follows that the approximate spectral densities at three frequencies can be obtained from <sup>15</sup>N R<sub>1</sub>, R<sub>2</sub>, and NOE measurements at a single field, by:

$$J(0.87\omega_H) = |\gamma_N/\gamma_H|(1 - NOE) R_1 / (5d^2) \quad (2.5.16)$$

$$J(\omega_N) = \frac{R_1 - 7(0.87/0.921)^2 d^2 J(0.87\omega_H)}{3(d^2 + c^2)} \quad (2.5.17)$$

$$J(0) = \frac{2R_2 - R_1 - 6(0.87)^2 d^2 J(0.87\omega_H)}{4(d^2 + c^2)}. \quad (2.5.18)$$

Alternatively, the high frequency contributions to the relaxation rates can be estimated using the above equations and subtracted from the measured relaxation rates to yield the so-called “reduced” relaxation rates<sup>52</sup>:

$$R_1' = R_1(1 - 1.249|\gamma_N/\gamma_H|(1 - NOE)) = 3(d^2 + c^2)J(\omega_N) \quad (2.5.19)$$

$$R_2' = R_2 - 1.079|\gamma_N/\gamma_H|R_1(1 - NOE) = 0.5(d^2 + c^2)(4J(0) + 3J(\omega_N)) \quad (2.5.20)$$

## Chapter 3: Methods

Since the discovery of nuclear magnetic resonance<sup>8,24</sup>, more than fifty years ago, the technique has found myriad applications. The ability to observe signals from individual nuclei of biomolecules in solution, which is their natural environment, makes NMR attractive for studying structure, dynamics, and interactions in proteins and nucleic acids. NMR of biomolecules would likely be impossible, however, without several technical developments—Fourier transform spectroscopy<sup>80</sup>, multi-dimensional spectroscopy<sup>81,82</sup>, pulse sequences making use of heteronuclei (pulse sequences making use of e.g.  $^{13}\text{C}$  or  $^{15}\text{N}$  as in all the experiments discussed below), and techniques for uniform and selective isotope labelling (to provide the spin  $\frac{1}{2}$  heteronuclei for these experiments)<sup>83</sup>, just to name a few. The possibility to extract distance information from proton NOESY spectra<sup>84</sup> led to the introduction of NMR as a technique for structural determination<sup>85</sup>, but since NMR protein structure determination is accomplished by a search of the conformational space accessible to the polypeptide chain for the structure which best complies with experimental constraints, it would be impossible for large molecules were it not for the development of efficient and robust computational algorithms and improvements over the past several decades in computer speed.

The development of experiments to measure spin-relaxation rates in proteins has led to the use of NMR to study motions in proteins. There exist experiments to measure auto-relaxation rates of  $^1\text{H}$ ,  $^2\text{H}$ , and  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei in proteins, with the purpose of obtaining information about motions of peptide planes ( $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^1\text{H}$ )

<sup>83,86,87</sup> and side chains (those already mentioned as well as <sup>2</sup>H <sup>88-91</sup>). Here I describe standard pulse-sequences to measure <sup>15</sup>N relaxation (<sup>15</sup>N R<sub>1</sub>, R<sub>2</sub> and the heteronuclear steady-state <sup>15</sup>N-<sup>1</sup>H NOE). These rates contain contributions from ultrafast (<0.1 picosecond) librational motion of N-H bonds which are fairly uniform throughout a protein, but are also sensitive to the anisotropic motions of the peptide-planes in the polypeptide backbone. Both the overall tumbling of the molecule and the fast (ps-ns) local motion of the peptide-planes contribute to these spin-relaxation rates. The deconvolution of these motions, and derivation of motional parameters from these rates will be explained and demonstrated in Chapters 4-6. Here I explain the technical details of the pulse sequences used for their measurement.

In section 3.3 I discuss the biochemical methods used to express, purify, and characterize the GB3 domain. As mentioned above, all of the NMR pulse sequences used in this study rely on uniform <sup>15</sup>N isotope labelling of the protein domain, and some of the experiments (see Chapters 3 and 7) require incorporation of <sup>2</sup>H nuclei in either the protein or the solvent. The proper purification of the protein domain at sufficient concentrations (>1 mM) for good signal-to-noise ratios in NMR spectra with reasonable experiment times is imperative for obtaining precise relaxation rates (see section 3.2). The accuracy and precision of the relaxation rates is pivotal for the derivation of accurate and precise structural and dynamic parameters from these data.

### 3.1 NMR Pulse Sequences for Measurement of $^{15}\text{N}$ Auto- and Cross-Relaxation Rates

#### **3.1.1 An Introduction to NMR Pulse Sequences**

In an NMR spectrometer, magnetization is excited (i.e. the populations of the spin states are perturbed from their equilibrium values) by radio-frequency pulses, which induce spin transitions; the magnetization is then allowed to evolve and is detected as a free-induction decay (FID). I will use the schematic diagram for an HSQC (Heteronuclear Single-Quantum Coherence) experiment shown in Figure 3.1.1 as an illustration of a “typical” 2D NMR experiment. Here and throughout, I write the angular momentum operators for  $^1\text{H}$  and  $^{15}\text{N}$  nuclei as  $H$  and  $N$ , and develop the formalism in terms of these operators. However, the equations and the experiments are generally applicable to any heteronuclear coupled spins- $1/2$  system, so that they could instead be written in terms of  $I$  and  $S$ , where  $I$  and  $S$  represent generalized angular momentum operators for any nuclei.

The three main steps of a 2D NMR pulse sequence are 1) excitation, 2) mixing, and 3) transfer and detection. These parts of the HSQC are indicated in Figure 3.1.1. In all steps the spin system evolves under the Zeeman Hamiltonian describing chemical shift and scalar coupling (J-coupling) interactions:

$$H = \omega_H H_z + \omega_N N_z + 2\pi J_{HN} H_z N_z \quad (3.1.1)$$

where  $\omega_H$  and  $\omega_N$  are the Larmor frequencies of the H and N nuclei respectively and  $J_{HN}$  is the scalar coupling constant between the H and N nuclei.

In the frame of reference which rotates about the direction of the external field (the z-axis in the laboratory frame) with angular frequency,  $\omega_x$ , each of the first two terms in the Hamiltonian can be written as:

$$H_{eff} = (\omega_x - \omega_{xrf}) X_z = \Omega_x X_z, \quad (3.1.2)$$

where  $X$  can represent either  $H$  or  $N$ ,  $\omega_{xrf}$  is the angular frequency of the rotating field for nucleus  $X$ , and  $\Omega_x$  is the chemical shift. This evolution is most conveniently described by a quantum mechanical “product operator” formalism, where the evolution of the density matrix describing the state of the two-spin system is followed in the basis of Cartesian components of the magnetization of the single spins  $H$  and  $N$ , where  $H$  and  $N$  refer to the angular momentum operators (in terms of the Pauli spin matrices) for the nuclei and of the two-spin order combinations  $H_i N_j$ , with  $i, j = x, y, z$ . Because the operators for the chemical shift evolution of spins  $H$  and  $N$  and the operator for the J-coupling evolution between spins  $H$  and  $N$  all commute with one another, the order in which the evolutions due to shift and coupling are considered is unimportant.

For spin-operator  $N$ , the chemical shift part of the Hamiltonian has the form  $\Omega_N N_z t$  where  $\Omega_N$  is the Larmor frequency of spin  $N$ . The chemical shift evolution during a period  $t$  is represented as:

$$\begin{aligned} N_z &\xrightarrow{\Omega_N N_z t} N_z \\ N_x &\xrightarrow{\Omega_N N_z t} N_x \cos(\Omega_N t) + N_y \sin(\Omega_N t) \\ N_y &\xrightarrow{\Omega_N N_z t} N_y \cos(\Omega_N t) - N_x \sin(\Omega_N t) \end{aligned} \quad (3.1.3)$$

with identical equations for the  $H_i$  components of magnetization (where  $i = x, y$ , or  $z$ ), substituting an  $H$  for  $N$  in 3.1.1 for proton evolution in terms of its chemical shift,



$\Omega_H$ . For a coupled two-spin system,  $H$  and  $N$ , the part of the Hamiltonian due to weak scalar coupling,  $J_{HN}$  is  $2\pi J_{HN} H_z N_z$  and the evolution of the  $H_i$  components of magnetization is described by:

$$\begin{aligned} H_z &\xrightarrow{2\pi J_{HN} H_z N_z t} H_z \\ H_x &\xrightarrow{2\pi J_{HN} H_z N_z t} H_x \cos(\pi J_{HN} t) + 2H_y N_z \sin(\pi J_{HN} t); \\ H_y &\xrightarrow{2\pi J_{HN} H_z N_z t} H_y \cos(\pi J_{HN} t) - 2H_x N_z \sin(\pi J_{HN} t) \end{aligned} \quad (3.1.4)$$

and similarly for the  $N_i$  components of magnetization:

$$\begin{aligned} N_z &\xrightarrow{2\pi J_{HN} H_z N_z t} N_z \\ N_x &\xrightarrow{2\pi J_{HN} H_z N_z t} N_x \cos(\pi J_{HN} t) + 2H_z N_y \sin(\pi J_{HN} t) \\ N_y &\xrightarrow{2\pi J_{HN} H_z N_z t} N_y \cos(\pi J_{HN} t) - 2H_z N_x \sin(\pi J_{HN} t) \end{aligned} \quad (3.1.5)$$

The evolution of the two-spin operators is described by:

$$\begin{aligned} 2H_z N_z &\xrightarrow{2\pi J_{HN} H_z N_z t} 2H_z N_z \\ 2H_y N_z &\xrightarrow{2\pi J_{HN} H_z N_z t} 2H_y N_z \cos(\pi J_{HN} t) - H_x \sin(\pi J_{HN} t) \\ 2H_x N_z &\xrightarrow{2\pi J_{HN} H_z N_z t} 2H_x N_z \cos(\pi J_{HN} t) + H_y \sin(\pi J_{HN} t) \end{aligned} \quad (3.1.6)$$

and:

$$\begin{aligned} 2H_z N_y &\xrightarrow{2\pi J_{HN} H_z N_z t} 2H_z N_y \cos(\pi J_{HN} t) - N_x \sin(\pi J_{HN} t) \\ 2H_z N_x &\xrightarrow{2\pi J_{HN} H_z N_z t} 2H_z N_x \cos(\pi J_{HN} t) + N_y \sin(\pi J_{HN} t) \end{aligned} \quad (3.1.7)$$

The Hamiltonian expression describing the radio-frequency pulses can be written as e.g.,  $H = \alpha N_x$  for an x pulse at the nitrogen resonance, or  $H = \alpha N_y$  for a y-pulse, where  $\alpha$  is the flip angle of the pulse. For example, the rotation for a pulse at the nitrogen resonance created by a  $\pm x$  pulse is given by:

$$\begin{aligned} N_z &\xrightarrow{\alpha N_{\pm x}} N_z \cos(\alpha) \mp N_y \sin(\alpha) \\ N_y &\xrightarrow{\alpha N_{\pm x}} N_y \cos(\alpha) \pm N_z \sin(\alpha). \\ N_x &\xrightarrow{\alpha N_{\pm x}} N_x \end{aligned} \quad (3.1.8)$$

For a  $\pm y$  pulse, the expressions are:

$$\begin{aligned} N_z &\xrightarrow{\alpha N_{\pm y}} N_z \cos(\alpha) \pm N_x \sin(\alpha) \\ N_y &\xrightarrow{\alpha N_{\pm y}} N_y \\ N_x &\xrightarrow{\alpha N_{\pm y}} N_x \cos(\alpha) \mp N_z \sin(\alpha) \end{aligned} \quad (3.1.9)$$

If the spin under study has a high sensitivity (the sensitivity of a nucleus is determined by its gyromagnetic ratio, and  $\gamma_{1H} = 2.675 \times 10^8 \text{ (T}\cdot\text{s)}^{-1}$  while  $\gamma_{13C} = 6.728 \times 10^7 \text{ (T}\cdot\text{s)}^{-1}$  and  $\gamma_{15N} = -2.712 \times 10^7 \text{ (T}\cdot\text{s)}^{-1}$ ) the excitation phase may consist of a simple, single radio-frequency pulse. Oftentimes the nuclei we want to study (e.g.  $^{15}\text{N}$  or  $^{13}\text{C}$  in proteins) are J-coupled to protons, and we can use this coupling to increase the signal from these “low sensitivity nuclei”. This type of excitation sequence is called an INEPT (Insensitive Nuclei Excitation via Polarization Transfer) sequence<sup>92</sup>. During the excitation period (between points **a** and **b** in Figure 3.1.1), the J-coupling between spins H and N creates  $2H_zN_y$  coherence (Eq. 3.1.5). The  $2H_zN_y$  coherence builds up with a rate proportional to the coupling constant,  $J_{HN}$ . Chemical shift evolution is refocused (caused to have zero net effect) on both channels H and N during the excitation step by sequences of the form  $\tau$ - $180^\circ_{N,H}$ - $\tau$  in both the H and N dimensions.

During the evolution period, (between points **b** and **c** in Figure 3.1.1) the magnetization is labeled for Fourier-transforming in the indirect dimension (the non-proton dimension, the dimension of spin N). This means that during this period, an additional time delay,  $t_1$  is introduced for the N spin, and this delay is varied from one repetition of the pulse sequence to the next, without summing the FIDs from the various repetitions. The acquisition of many data sets with different values of  $t_1$  along

with the acquisition time (called  $t_2$ ) leads to compilation of a two-dimensional signal surface that is a function of variables ( $t_1, t_2$ ) and can be Fourier-transformed into a bivariate function of ( $\omega_1, \omega_2$ ). This allows the signal originating from the spin-system to be separated into a 2D map with one characteristic signal from each unique chemical environment of the pair of nuclei H and N. During the evolution phase of the experiment, it is common to employ a  $180^\circ$  pulse on spin H (see Figure 3.1.1) in order to cancel the effect of J-coupling during the delay  $t_1$ . Alternatively, this decoupling can be performed by a train of pulses (labeled GARP in a pulse sequence, GARP stands for Globally optimized Alternating phase Rectangular Pulse decoupling).

The mixing part of the NMR experiment involves transfer of magnetization back to the high sensitivity nucleus for detection and is followed by the detection itself. The transfer is accomplished by a backwards form of the INEPT sequence used during the excitation. Decoupling pulses are also applied during the detection period to ensure that the detected signal is modulated only by the chemical shift of spin H and not by the J-coupling constant,  $J_{\text{NH}}$ .



### 3.1.2 Longitudinal Relaxation, $R_1$

In the absence of a magnetic field, the nuclear spin polarizations of molecules in solution are uniformly distributed in space—this means that the net nuclear magnetic moment of the sample is small. If a magnetic field is applied, thermal motions of the molecule in the sample result in a net magnetic anisotropy of the system, due to the extremely small energetic advantage of alignment of the spin with the magnetic field ( $\hbar\omega_H$  at 600 MHz is  $< 4 \cdot 10^{-25}$  Joules, which is  $\sim 10,000$  times smaller than  $k_B T$  at 20°C). If this magnetic moment is rotated into the plane perpendicular to the applied field by means of an rf pulse, the net nuclear spin magnetization in the z-direction will relax back to its equilibrium value according to the relationship:

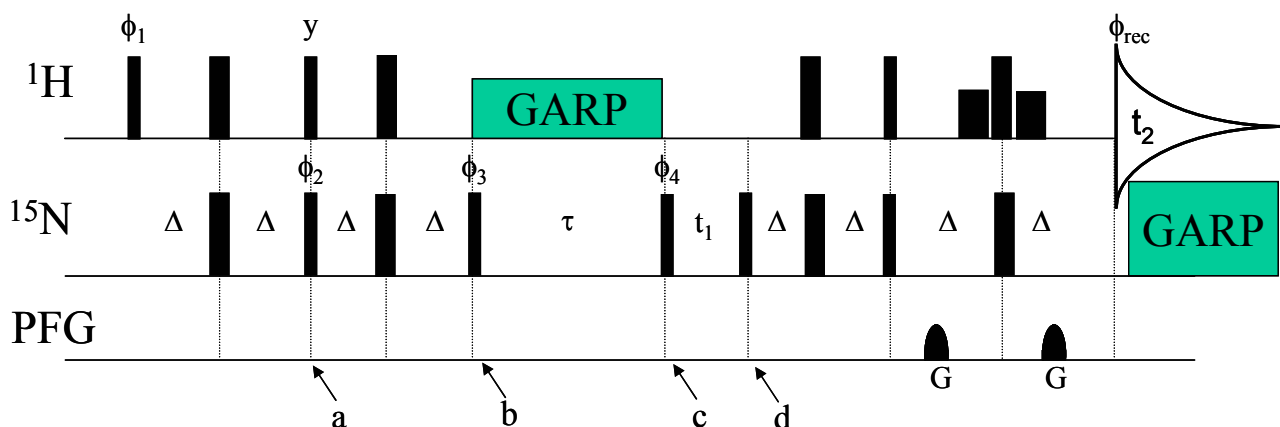
$$M_z(t) = M_0(1 - e^{-R_1 t}),$$

where the rate constant,  $R_1$ , is called the longitudinal relaxation rate of the nuclear magnetic moment or the spin-lattice relaxation rate.

The pulse sequence used to measure  $R_1$  is shown in Figure 3.1.3. Here the magnetization in the z-direction is inverted with respect to its equilibrium state then allowed to relax. The excitation phase of this experiment has two steps: the first one, up to point **a** in the sequence, excites two-spin order antiphase coherence  $2N_yH_z$  via the J-coupling evolution during the interval  $2\Delta$  (according to Eq.3.1.7); the second part of the excitation phase, from point **a** to point **b**, transforms the antiphase nitrogen magnetization into in-phase magnetization ( $2H_zN_y \rightarrow N_x$ ), according to Eq.3.1.7. As in the HSQC, the delay  $\Delta$  is chosen such that  $2\Delta = 1/(2J_{NH})$  in order to cancel the cosine-modulated components. For the  $^{15}\text{N}$ - $^1\text{H}$  J-coupling in the backbone amide bonds of

proteins,  $J_{\text{NH}}$  is approximately 94 Hz, hence  $\Delta = 1/4J_{\text{NH}} \sim 2.7$  ms, which allows for short transfer periods without significant losses of magnetization through transverse relaxation. The nitrogen single-quantum coherence prior to point **b** ( $N_x$ ) is transformed into zero-quantum coherence ( $N_z$ ) using the  $90^\circ$  pulse at point **b** and allowed to relax towards equilibrium with relaxation rate  $R_1$ . The resulting magnetization is flipped back into the xy plane at the end of the relaxation interval and subsequently labeled by chemical shift evolution during  $t_1$ . Composite-pulse decoupling is applied on the  $^1\text{H}$  spins during the relaxation delay (between points **b** and **c**, labeled GARP) to minimize systematic contributions from the time-dependent effects of dipolar  $^{15}\text{N}$ - $\{^1\text{H}\}$  cross relaxation (see section 3.1.4) and of cross-correlation between dipolar and chemical-shift anisotropy relaxation mechanisms (see Chapter 4). The following reverse-INEPT period transform nitrogen single-quantum in-phase coherence into proton single-quantum coherence prior to detection.

**Figure 3.1.3. Pulse sequence for measurement of longitudinal relaxation rate,  $R_1$ .** The phase cycling is  $\phi_1=x,-x$ ;  $\phi_2=2(x),2(-x)$ ;  $\phi_3=4(x),4(-x)$ ;  $\phi_4=8(x),8(-x)$ ; and the receiver  $\phi_{\text{rec}}=2(x,-x), 2(-x,x), 2(-x,x), 2(x,-x)$ . The ideal value for  $\Delta$  is such that  $2\Delta=1/(2J_{\text{IS}})$ . The phase cycle for  $\phi_4$  is such that both cosine and sine modulations of the chemical shift of  $^{15}\text{N}$  are detected during the  $t_1$  delay via time-proportional increments (States-TPPI). GARP stands for Globally Optimized Alternating Phase Rectangular Pulse Decoupling. The two low-power  $90^\circ$  pulses flanking the last nonselective  $180^\circ$  pulse on protons (as part of the water suppression element) are 1ms long.



### 3.1.3 Transverse Relaxation, $R_2$

If the net magnetic moment of the sample at thermal equilibrium in a magnetic field (as discussed above) is rotated through an angle of  $\pi/2$  about an axis perpendicular to the field by application of a radiofrequency pulse, the net spin polarization is in the plane transverse to the magnetic field of the spectrometer. This transverse magnetic moment will begin to relax towards zero as precessing spins lose coherence with each other. This relaxation is given by the relationship:

$$M_y(t) = -M_0 \cos(\omega_0 t) e^{-tR_2}, \text{ and } M_x(t) = M_0 \sin(\omega_0 t) e^{-tR_2},$$

where x and y are the axes perpendicular to the applied field (which is defined as in the z direction). The spin relaxation rate,  $R_2$ , is called the transverse spin relaxation rate or the spin-spin relaxation rate.

The pulse sequence for  $R_2$  measurement (Fig. 3.1.4) is very similar to the one used for measuring  $R_1$ . Nitrogen single quantum in-phase magnetization is created at time point **a** in order to observe the decay of this coherence as a function of the evolution delay  $4n\tau$  (adaptation of the spin-echo experiment with the Carr-Purcell-Meiboom-Gill (CPMG) sequence<sup>93,94</sup>). During the evolution period, each  $180^\circ$  pulse about the x-axis rotates the equilibrium magnetization into the xy plane. The different effective magnetic fields felt by all the different nuclei in the sample cause the magnetization vectors from different nuclei to “fan out” i.e. they rotate with slightly different Larmor frequencies depending on the strength of the effective field where they are located (i.e. on their chemical shift as well as any field inhomogeneity). After some arbitrary time, an  $180^\circ$  pulse about the y axis reflects all the vectors in the xy plane and they continue to rotate, after a further time of equal duration to the first





### 3.1.4 Heteronuclear $^{15}\text{N}$ - $\{^1\text{H}\}$ NOE

The heteronuclear steady-state nuclear Overhauser effect (NOE) experiment is a measurement of the heteronuclear NOE enhancement, which is related to the cross relaxation rate constant,  $\sigma_{\text{IS}}$ . The NOE enhancement for a two spin system is most frequently measured using the steady-state NOE difference experiment. In this experiment, two spectra are recorded and the ratio of the peaks in the two spectra is related to the NOE effect. In one spectrum, the  $H$  spin is saturated for a period of time long enough to establish the NOE effect on the  $N$  spin. Then a reverse INEPT transfers magnetization back to spin  $H$  for detection and the free-induction decay (FID) is recorded. The intensity of the peaks in this spectrum will be proportional to the NOE effect. In the other spectrum, the signal is recorded without proton saturation. The intensity of the peaks in this spectrum will be proportional to the equilibrium values (no NOE effect), and the NOE can then be calculated from the ratio of the intensities of the peaks in the first spectrum to the intensities of the peaks in the second spectrum.

Water suppression in the NOE experiments was accomplished using the flip-back scheme<sup>95</sup> to avoid problems due to amide proton-solvent exchange with a recycling delay of 4-5 s (magnetization transfer from water to amide protons, either via the dipolar interaction or hydrogen exchange, prohibits the use of water presaturation, as it can cause amide protons to relax to their thermal equilibrium value faster than their inherent  $R_1$ ).



**Table 2.2.1. Relaxation delays for all experiments.**

| Frequency            | Measured rate                                     | Relaxation delay  |
|----------------------|---|---|
| 400 MHz ( 9.4 Tesla) | R <sub>1</sub>                                    | 432, 432, 4, 432, 432, 4 ms   |
|                      | R <sub>2</sub>                                    | 280, 280, 8, 280, 280, 8 ms   |
|                      | NOE   | 4.5 s   |
|                      | $\eta_{xy}$                                       | 0, 21.27, 31.91, 42.53, 53.19 ms  |
|                      | $\eta_z$  | 0, 100, 200, 250, 300, 350 ms   |
|                      |   |   |
| 500 MHz ( Tesla)     | R <sub>1</sub>                                    | 504, 504, 4, 504, 504, 4 ms   |
|                      | R <sub>2</sub>                                    | 248, 248, 8, 248, 248, 8 ms   |
|                      | NOE   | 4.69 s  |
|                      | $\eta_{xy}$                                       | 0, 31.91, 26.59, 42.55, 53.19 ms  |
|                      | $\eta_z$  | 0, 100, 200, 250, 300, 350 ms   |
|                      | R <sub>1</sub> ( <sup>15</sup> N- <sup>2</sup> H) | 0.15, 105.84, 209.13, 295.21, 398.51 (x2), 605.1, 794.48 and 1001.07 ms |
|                      | R <sub>2</sub> ( <sup>15</sup> N- <sup>2</sup> H) | 3.74, 103.10, 198.78 (x2), 265.02, 323.90, 397.50 and 449.02 ms         |
|                      |   |   |
| 600 MHz ( Tesla)     | R <sub>1</sub>                                    | 440, 440, 4, 440, 440 ms  |
|                      | R <sub>2</sub>                                    | 264, 264, 8, 264, 264 ms  |
|                      | NOE   | 5 s   |
|                      | $\eta_{xy}$                                       | 26.59, 31.91, 37.23, 42.55, 53.19 ms                                    |
|                      | $\eta_z$  | 0, 100, 150, 200, 300, 400 ms   |
|                      | $\eta_z$ (perdeuterated GB3)                      | 0, 100, 150, 200, 300, 400 ms   |
|                      | R <sub>1</sub> $\rho$ (repetition delay)          | 1, 4, 8 ms  |
|                      |   |   |
| 700 MHz ( Tesla)     | R <sub>1</sub>                                    | 620, 620, 4, 620, 620, 4 ms   |
|                      | R <sub>2</sub>                                    | 264, 264, 8, 264, 264, 8 ms   |
|                      | NOE   | 4 s   |
|                      | $\eta_{xy}$                                       | 26.59, 31.91, 37.23, 42.55, 53.19 ms                                    |
|                      |   |   |
| 800 MHz ( Tesla)     | R <sub>1</sub>                                    | 672, 672, 4, 672, 672, 4 ms   |
|                      | R <sub>2</sub>                                    | 248, 248, 8, 248, 248, 8 ms   |
|                      | NOE   | 4.7 s   |
|                      | $\eta_{xy}$                                       | 26.59, 31.91, 37.23, 42.55, 53.19 ms                                    |

Spectral peak intensities were extracted using software written in house (AUTOPICK) in the Matlab programming environment. The position of maximum

intensity for each peak is found by iteratively moving the peak position from some user-supplied starting position for as many iterations as the intensity increases in either of the two-dimensions of the spectrum. The intensity of the peak at the maximum is determined using three-point parabolic interpolation. As a control, peak volumes for all planes in the 600 MHz  $R_1$  experiment were extracted using the PEAKINT subroutine of the XEASY software. The peak heights and peak volumes yielded equivalent values of the relaxation parameters from the exponential fitting procedure; however, slightly better fits (smaller  $\chi^2$ ) and lower uncertainties were obtained for the peak-height analysis, therefore peak heights are used in all the analyses herein. The better fits of the peak intensities are possibly due to problems in the fitting of the lineshape in the peak-volume determination which are not present in the height determination, or due to spectral artifacts which influence the peak volumes but do not affect the peak intensities.

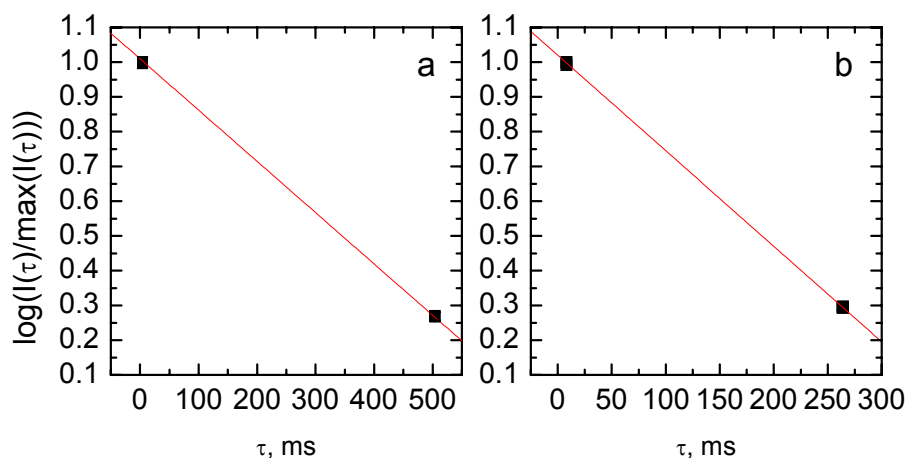
The relaxation rates ( $R_1$  and  $R_2$ ) were obtained by least-square fitting of peak intensities in the corresponding series of 2D spectra to a mono-exponential decay. The fitting function has the form:

$$I_j(\tau) = I_j(0)e^{-R_j\tau} \quad , \quad (3.2.1)$$

where  $I_j(t)$  is the intensity of a particular peak,  $j$ , as a function of relaxation time,  $\tau$ ;  $R_j$  is the fit relaxation rate ( $R_1$  or  $R_2$ ) constant for residue  $j$ , and  $I_j(0)$  is the intensity of peak  $j$  at time  $\tau = 0$  and was fit simultaneously along with  $R$ . The fit was performed using a Nelder-Mead simplex-based multivariable  $\chi^2$  minimization. The heteronuclear NOE values were obtained from the ratio of peak intensities in the NOE and NONOE experiments. The exponential decays for residue Tyr3 from the

600 MHz  $R_1$  and  $R_2$  experiments are shown in Figure 3.2.1. The residuals from the experimental points and a fitting curve were used to estimate the width of the assumed-normal distribution of experimental points about the best fit curve. Random values of this distribution (generated by Monte-Carlo simulation of 500 synthetic data sets per residue) were then used to generate synthetic data sets, each with approximately the same residuals as the experimental data. These data were then fit to an exponential curve, to yield an ensemble of values for the decay rate. The uncertainty in the rate is the standard deviation of this ensemble. The obtained uncertainties depend on the input uncertainties in peak intensities—the uncertainties in peak intensities were calculated in two different ways (see below).

**Figure 2.2.1. Linear fit of intensity decay on natural log scale for  $R_1$  and  $R_2$  experiments for Tyr3 in GB3. Panel a shows  $R_1$  fit, where two data points were collected at 4 ms and 4 data points at 504 ms. Panel b shows the  $R_2$  fit, where there are 2 data points at 8 ms and 4 points at 264 ms. Errors in intensities are smaller than data points in both fits.**



The means and standard deviations of the signal in several sections of the spectra that did not contain cross peaks or other signals (such as the water resonance)

were calculated to provide an estimate of the baseline offset and rms baseline noise level ( $\sigma_i$ ). In all cases the mean was approximately zero, and less than the standard deviation, indicating that there was no appreciable baseline offset. The value of  $\sigma_i$  in all experiments was inversely correlated with the relaxation time consistent with the observation of Skelton et al.<sup>97</sup>. In experiments where it was possible, an estimate of the uncertainty in peak intensities was also obtained by comparison of intensities of peaks in duplicate (or quadruplicate in some cases) spectra using the method of Skelton et al.<sup>97</sup>. The values obtained here, for the standard deviation in the peak intensities,  $\sigma_r$ , are between 1.5 and 10 times greater than the corresponding  $\sigma_i$  values, and the  $\sigma_r$  values are similarly inversely proportional to the relaxation time (this is in general agreement to previous observations for sensitivity enhanced relaxation rate measurements<sup>97</sup>, although here we did not examine the detailed time dependence of  $\sigma_r$  and  $\sigma_i$ ). For all experiments, the use of the  $\sigma_r$  values for the uncertainty in peak intensities resulted in uncertainties in the fit relaxation rates between 0.5%-2% (using a Monte-Carlo simulation of the peak intensities, see above) whereas using the  $\sigma_i$  values resulted in uncertainties between 0.05% and 0.7%. As a test of how well these estimates represent the actual reproducibility of the relaxation rates, duplicate  $R_1$  data sets were acquired at 600 MHz. The  $R_1$  values from the two data sets are extremely similar, with  $R_1$  rates for all residues being within 3%, indicating that our estimates of the uncertainty in the rates using the  $\sigma_r$  estimate of the noise in peak heights is reasonable.

The uncertainties in the peak heights in the NOE experiments were estimated using the signal-to-noise ratio from integration of an area of the spectra (in the NOE

and NONOE experiments) containing no cross peaks. These errors were then propagated into the error in the NOE cross relaxation rate.

### **3.3 Protein Expression, Purification, and Characterization**

The GB3 sample on which relaxation rates at 400-800 MHz were measured was a gift from the National Institutes of Health. This sample contained 1.8 mM uniformly  $^{15}\text{N}$  enriched GB3 in 265 mL of 30 mM phosphate buffer at pH 5.8 and approximately 9% (by volume)  $\text{D}_2\text{O}$ . The 5.8 mM  $^{15}\text{N}$  labeled samples used for the direct nitrogen detection relaxation rate measurements as well as the perdeuterated sample used for comparison of longitudinal cross-correlation rates, were expressed and purified in the lab. The plasmid for these samples, which was also provided by NIH, was cloned into *E.coli* HMS174( $\lambda$ DE3) cells. Starter cultures were grown for 8 hours at 37°C to an optical density value at 600 nm,  $\text{OD}_{600}$ , greater than 0.5 using isolated bacterial colonies from plated cells. The cell culture was grown in auto-inducing ZYP-5052 medium (recipe from Dr. William Studier, Brookhaven National Laboratory) with  $(\text{NH}_4)_2\text{SO}_4$  replaced with  $^{15}\text{NH}_4\text{Cl}$  and  $\text{Na}_2\text{SO}_4$  so that the only Nitrogen source available to the bacteria was  $^{15}\text{N}$  (isotope enrichment 99%). The cultures were also grown with 100mg ampicillin per liter of culture, as the *E.coli* cells are genetically modified to be resistant to this antibiotic. This culture was grown overnight (to an  $\text{OD}_{600}>1$ ) at 37°C in a shaker incubator which agitates at 200 rpm.

The cells were harvested and resuspended into 1xPBS (~5 ml per gram of cell paste), then this suspension was heated at 80° C for 15 min (vortexed 2-3 times during heating), then cooled on ice for 15 min. This processes served to “heat shock” the cells and causes efficient cell lysis. The  $T_m$  of GB3 is >86°C, so the protein is not

unfolded by the heating. The lysed cell suspension was then centrifuged at 16,000 rpm in a 45Ti rotor for 30 min. The supernatant was then filtered and dialyzed overnight (at 4°C in 3.5kDa cutoff dialysis tubing) against 1.5 liters of 50 mM NaPO<sub>4</sub>, 0.5 M NaCl, and 0.1% sodium azide, at pH 5.5.

The dialyzed sample was then concentrated in a 5 kDa cutoff concentrator unit in a table-top centrifuge to a volume of 8-10 ml. The GB3 sample was purified by size exclusion chromatography. For purification, the 8-10 ml of sample was divided over 4-5 separate but identical column runs of 2 ml sample (filtered through 0.45 µm syringe filter) each on a Pharmacia Superdex 16/60 75 Hi-Load Prep column. The column was equilibrated with 3-4 column volumes of 50 mM NaPO<sub>4</sub>, 0.5 M NaCl, and 0.1% azide, pH 5.5 buffer. The FPLC was run at 0.3 ml/min and fractions of 5-10 ml were collected. The purified protein was examined by gel electrophoresis, and concentrated in a 5 kDa cutoff concentrator unit in a table-top centrifuge, then exchanged into 30 mM phosphate buffer with pH 5.8. The protein sample concentration was determined using absorbance at 280 nm (the extinction coefficient,  $\epsilon$ , of GB3 at 280 nm is 8250 M<sup>-1</sup>cm<sup>-1</sup>). Approximately 9% (by volume) D<sub>2</sub>O was added to NMR samples.



## Chapter 4: CSA/dipolar Cross-Correlated Relaxation Rates

### 4.1 CSA/Dipolar Cross-Correlated Relaxation

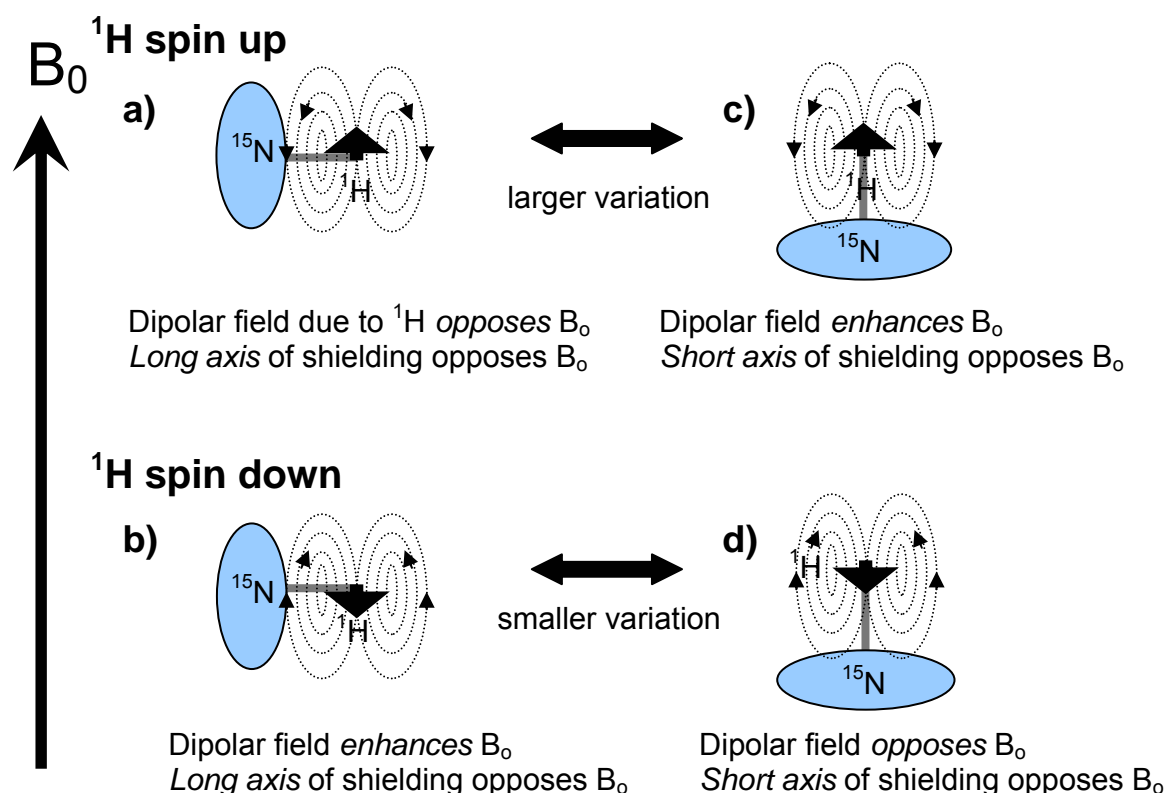
$^{15}\text{N}$ - $\{^1\text{H}\}$  CSA/dipolar cross-correlation was introduced in Chapter 2 where relaxation rates arising from cross-commutators ( $[[Q, T_m^i], T_m^j]$  and  $[[Q, T_m^j], T_m^i]$  where  $i \neq j$  in Eq. 2.3.1) in the equation for evolution of an observable under the master equation were discussed and the expressions for these rates in terms of spectral density functions were given in Eqs. 2.5.7-2.5.8. In Chapter 3 cross-correlated relaxation effects were discussed as a possible complication to the measurement of  $^{15}\text{N}$  autorelaxation rates. A physical explanation of the mechanism of CSA/dipolar cross-correlation is still needed; Figure 4.1.1 is a cartoon qualitatively depicting CSA/dipolar cross-correlation in a fictional 2D molecule. Here 2D nuclear spin,  $^{15}\text{N}$ , in an external magnetic field,  $B_0$ , and experiencing a through-space dipole-dipole interaction with a covalently bound  $^1\text{H}$  spin is depicted. The chemical shielding of the nitrogen nucleus is represented as an ellipse, where the long axis of the ellipse represents the most-shielded component (low field) of the chemical shielding tensor and the short axis of the ellipse represents the least-shielded (high field) component. The spin state of the  $^1\text{H}$  nucleus is represented by a thick arrow with spin state  $H_z = +\frac{1}{2}$  an upwards pointing arrow and  $H_z = -\frac{1}{2}$  a downwards pointing arrow. The dipolar field generated by the  $^1\text{H}$  nucleus is represented by dotted field lines. From the left panels (top and bottom) it is clear that if the nuclear spin of the  $^1\text{H}$  nucleus is aligned with the field (corresponding to spin state  $H_z = +\frac{1}{2}$ ), the magnetic field at the  $^{15}\text{N}$  spin ( $B_N$ ) due to the dipolar coupling and the field anisotropy of the shielding tensor tend to reinforce each other. If the nuclear spin of the  $^1\text{H}$  nucleus is aligned anti-parallel to

the external field (corresponding to spin state  $H_z = -1/2$ ), the dipolar and shielding anisotropy terms tend to oppose each other. This causes the magnetic field felt by the subpopulation of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spins in the spin-up ( $H_z = +1/2$ ) position to be different than the magnetic field felt by the subpopulation of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spins in the spin-down position ( $H_z = -1/2$ ) which results in two signals for each  $^{15}\text{N}$  nucleus (a  $^{15}\text{N}$  doublet) in a coupled 2D NMR spectrum. As discussed in Chapter 2,  $^{15}\text{N}$  spin relaxation is caused by time-dependent fluctuations in the magnetic field felt by the  $^{15}\text{N}$  nucleus. These fluctuations are caused by the tumbling of the molecule in solution and local motions of the peptide planes and N-H bond. This is shown in Figure 4.1.1, where the left and right sides of the figures represent two different orientational states of the molecule with respect to the magnetic field. The variations/fluctuations in  $B_N$  due to the tumbling of the molecule are represented by  $\Delta B_N$  in the caption to Figure 4.1.1. The magnitude of these fluctuations will be larger for the subpopulation of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spins in the spin-up position than for the subpopulation of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spins in the spin-down position. This causes the subpopulation of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spins in the spin-up position to relax faster than the subpopulation of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spins in the spin-down position. Thus the two components of the  $^{15}\text{N}$  doublet have different relaxation rates and different linewidths.

This differential relaxation is called  $^{15}\text{N}\{-^1\text{H}\}$  CSA/dipolar relaxation interference or CSA/dipolar cross-correlation and the rates resulting from this relaxation interference are called  $^{15}\text{N}\{-^1\text{H}\}$  CSA/dipolar cross-correlation rates (CCRs). The  $^{15}\text{N}$  CSA/dipolar CCRs are of particular interest because they allow

measurement of the “projection” (through the Legendre Polynomial) of the  $^{15}\text{N}$ - $^1\text{H}$  internuclear vector on the  $^{15}\text{N}$  CSA <sup>98,99,100</sup>.

**Figure 4.1.1** Schematic illustration of the effects of  $^{15}\text{N}$ - $^1\text{H}$  CSA/dipolar cross-correlation in a 2D model molecule. The anisotropic chemical shielding of the  $^{15}\text{N}$  nucleus is represented by an ellipse where the long axis of the ellipse corresponds to the most shielded component of the chemical shift tensor and the short axis corresponds to the less shielded component. The left-hand panels represent one 2D molecular orientation of the  $^{15}\text{N}$ - $^1\text{H}$  spin system with the  $^1\text{H}$  spin in the spin up (spin state  $H_z=+\frac{1}{2}$  (top panel)) and spin down ( $H_z=-\frac{1}{2}$  (bottom panel)) and the right-hand panels represent another 2D molecular orientation with the up (down)  $^1\text{H}$  spin states also in the top (bottom) panels. In a), the total magnetic field at the location of the  $^{15}\text{N}$  nucleus is given by  $B_N=B_0(1-\sigma_{\text{iso}})-B_d(\Delta\sigma/2)$  where  $B_d$  is the contribution to the magnetic field due to the dipolar interaction with  $^1\text{H}$  and  $\sigma_{\text{iso}}$  and  $\Delta\sigma$  are the isotropic chemical shielding and the CSA (defined such that both are positive quantities). In b)  $B_N=B_0(1-\sigma_{\text{iso}})+B_d(\Delta\sigma/2)$ , in c)  $B_N=B_0(1-\sigma_{\text{iso}})+B_d(\Delta\sigma/2)$  and in d)  $B_N=B_0(1-\sigma_{\text{iso}})-B_d(\Delta\sigma/2)$ . The variation in the field  $B_N$  as a)→c) (for  $^1\text{H}$  spin up)  $\Delta B_N=2B_d+\Delta\sigma$ , is larger than the variation in  $B_N$  as b)→d) (for  $^1\text{H}$  spin down)  $\Delta B_N=-2B_d+\Delta\sigma$ , therefore the relaxation rate of the population of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spin up relaxes faster than the population of  $^{15}\text{N}$  spins bound to  $^1\text{H}$  spin down.



Measurements of  $^{15}\text{N}$  CSA/dipolar cross-correlation rates can be used to characterize the overall and internal motions in proteins and nucleic acids <sup>43,51,101,102</sup>,

<sup>103</sup>) and to determine the magnitude and orientation of <sup>15</sup>N chemical shift tensors in proteins<sup>50,51,56</sup>. In Chapter 6 of this document, I describe how we use  $\eta_{xy}$  (the transverse <sup>15</sup>N CSA/dipolar CCR) and  $\eta_z$  (the longitudinal CCR discussed in section 4.4) measured in GB3 in combination with autorelaxation rates  $R_1$  and  $R_2$  to identify residues undergoing chemical exchange motions. In Chapter 7, I compare rotational diffusion properties of GB3 derived from  $\eta_z$  and  $\eta_{xy}$  to properties derived from autorelaxation rates, and use  $\eta_z$  and  $\eta_{xy}$  (in combination with other rates) to determine site-specific chemical shielding anisotropies in GB3.

## **4.2 Measurement of CSA/Dipolar CCRs from Coupled HSQC Spectra**

### **4.2.1 Direct versus Indirect Methods for CCR measurement**

As discussed above, there is great interest in the measurement of CSA/dipolar CCRs in proteins, hence many pulse sequences have been suggested for the measurement of these CCRs<sup>43,99,104-108</sup>. In an NMR experiment, the transverse <sup>15</sup>N-<sup>1</sup>H CSA/dipolar CCR,  $\eta_{xy}$  given in terms of the spectral density functions in Eq. 2.5.7-2.5.10, describes how anti-phase two-spin order magnetization (e.g.  $N_y H_z$ ) builds up from nitrogen single-spin order magnetization (e.g.  $N_y$ ). The equations describing the decay and build-up of transverse components of magnetization for an isolated <sup>15</sup>N-<sup>1</sup>H spin system in terms of the product operators describing magnetization in an NMR experiment as introduced in Chapter 3 (Eqs. 3.12-3.13) are:

$$\frac{\partial}{\partial t} \begin{pmatrix} \langle N_y(t) \rangle \\ \langle 2N_y H_z(t) \rangle \end{pmatrix} = - \begin{pmatrix} R_{2,N} & \eta_{xy,N} \\ \eta_{xy,N} & R_{2,2HN} \end{pmatrix} \cdot \begin{pmatrix} \langle N_y(t) \rangle \\ \langle 2N_y H_z(t) \rangle \end{pmatrix}, \quad (4.2.1)$$

where  $R_{2,N}$  and  $R_{2,2HN}$  are the transverse auto-relaxation rates of nitrogen and antiphase two-spin order magnetization, and  $\eta_{xy,N}$  is the transverse CCR of the <sup>15</sup>N-<sup>1</sup>H

dipolar interaction with the  $^{15}\text{N}$  CSA. If we consider separately the projection of any nitrogen spin operator ( $N_y$ ) onto the subspace  $H_z=+1/2$  and  $H_z=-1/2$ , then:

$$N_y^{1H(up)} = N_y \left( \frac{1}{2} + H_z \right) \quad \text{and} \quad N_y^{1H(down)} = N_y \left( \frac{1}{2} - H_z \right). \quad (4.2.2)$$

Where  $N_y^{1H(up)}$  is the projection of  $N_y$  onto the  $H_z=+1/2$  subspace and  $N_y^{1H(down)}$  is the projection of  $N_y$  onto the  $H_z=-1/2$  subspace. It is computationally more convenient to use the operators in Eq. 4.2.1, which in terms of these projections are given by:

$$N_y = N_y^{1H(up)} + N_y^{1H(down)} \quad (4.2.3)$$

$$2N_y H_z = N_y^{1H(up)} - N_y^{1H(down)}. \quad (4.2.4)$$

The existing approaches to measuring rates of CSA/dipolar cross-correlated relaxation (CCR) can be divided in two classes<sup>109</sup>: J-resolved (or “direct”) and quantitative (or “indirect”) experiments. Most of these sequences can be characterized as indirect methods of measurement in that they measure the rate of conversion of one coherence into another via (at least) two separate experiments (e.g.<sup>43,104</sup>). One experiment, denoted “A” is needed to measure the build up of the selected coherence (e.g.  $N_y$  from  $2N_y H_z$ ) and another experiment, “B” to measure the autorelaxation of the initial component (e.g.  $2N_y H_z$  from  $2N_y H_z$ ). Small differences in experimental conditions or pulses between the A and B experiments can lead to deviations in the signal ratio between the two experiments and thus cause inaccurate measurements of the CCR. The published pulse sequences differ between experiments A and B in the number of pulses<sup>43</sup> or in the order in which the selection elements are applied<sup>104</sup>. Imperfections in the parameters of these pulse sequences can lead to incomplete suppression of cross-correlated relaxation before and after the mixing period<sup>109</sup> and

can introduce deviations in the signal ratio between the two experiments that are difficult to quantify, as there is no direct control of the magnetization pathways. Recently an experimental scheme was proposed which attempts to completely balance differences in the evolution of the generated and detected coherences using the indirect approach<sup>108</sup>; this scheme requires four separate experiments at each time point in the CCR decay (two for measurement and two for normalization) to ensure that all possible deviations are balanced.

As explained schematically with Figure 4.1.1, the difference in linewidths of the up-field and down-field components of the  $^1\text{H}$ - $^{15}\text{N}$  scalar-coupled doublet is a direct result of CSA/dipolar cross-correlation relaxation. Therefore, it is straightforward to measure this rate constant from the time evolution of the ratio of the intensities of these components, in a direct fashion. In principle, the decay of the ratio of intensities is mono-exponential, the components of the nitrogen doublet are given by:

$$\sigma_{up} = \langle N_y^{1H(up)} \rangle = \left\langle \frac{1}{2} (N_y + 2N_y H_z) \right\rangle \quad (4.2.5)$$

$$\sigma_{dn} = \langle -N_y^{1H(dn)} \rangle = \left\langle -\frac{1}{2} (N_y - 2N_y H_z) \right\rangle. \quad (4.2.6)$$

Using Eq. 4.1.2, the time derivatives of the up-field and down-field components are given by:

$$\frac{d}{dt} \sigma_{up} = -i \frac{J_{NH}}{2} \langle N_y^{1H(up)} \rangle - (R_2^{av} + \eta_{xy}) \langle N_y^{1H(up)} \rangle - (R_2^{1/2 dif}) \langle N_y^{1H(dn)} \rangle \quad (4.2.7)$$

$$\frac{d}{dt} \sigma_{dn} = -i \frac{J_{NH}}{2} \langle N_y^{1H(dn)} \rangle + (R_2^{av} - \eta_{xy}) \langle N_y^{1H(dn)} \rangle + (R_2^{1/2 dif}) \langle N_y^{1H(up)} \rangle \quad (4.2.8)$$

where  $J_{NH}$  is the one-bond scalar coupling and  $R_2^{av}$  and  $R_2^{\frac{1}{2}dif}$  are the average and half-difference of the relaxation rates for nitrogen in-phase ( $N_y$ ) and anti-phase

( $2N_yH_z$ ) coherences. In proteins, the half-difference,  $R_2^{\frac{1}{2}dif} = \frac{R_{2,N} - R_{2,2N_yH_z}}{2}$ , of

these rates is always smaller than the average,  $R_2^{av} = \frac{R_{2,N} + R_{2,2N_yH_z}}{2}$ . If the relaxation

delays are chosen such that they are multiples of  $\frac{1}{4J_{NH}}$ , the two terms  $R_{2,N}$  and

$R_{2,2N_yH_z}$  are averaged exactly<sup>99</sup> so that the half-difference is zero for the period of the relaxation delay. Therefore the transverse relaxation rates of the two doublet components are given by:

$$\frac{d}{dt}\sigma^{up} = R_2^{av} + \eta_{xy} \text{ and } \frac{d}{dt}\sigma^{dn} = R_2^{av} - \eta_{xy} \quad (4.2.9)$$

Therefore the decay of the ratio of the volume of the upfield peak to the volume of the downfield peak is given by:

$$\frac{d}{dt} \left( \frac{\sigma^{up}}{\sigma^{dn}} \right) = \frac{e^{-(R_2^{av} + \eta_{xy})t}}{e^{-(R_2^{av} - \eta_{xy})t}} = e^{-2\eta_{xy}t}. \quad (4.2.10)$$

If instead the ratio of the intensities is taken, the expression will contain a time-independent pre-factor,  $A$ :

$$\frac{d}{dt} \left( \frac{S^{up}}{S^{dn}} \right) = \frac{A^{up} e^{-(R_2^{av} + \eta_{xy})t}}{A^{dn} e^{-(R_2^{av} - \eta_{xy})t}} = A e^{-2\eta_{xy}t} \quad (4.2.11)$$

where  $S^{up}$  and  $S^{dn}$  are the intensities of the upfield and downfield peaks, and  $A^{up}$  and  $A^{dn}$  are factors related to their lineshapes, with  $A = A^{up}/A^{dn}$ . Measurement of the CCR, then, in principle requires only one coupled HSQC-type experiment at each time

point in the CCR decay, and on small proteins or proteins with particularly good peak separation this is the simplest and most straightforward way to measure the CCR. A pulse sequence for this type of measurement is shown in Figure 4.2.1 (called in-phase, IP). The relaxation of interest takes place during the constant-time evolution period  $2\Delta$ . Protons are not decoupled during the  $^{15}\text{N}$  evolution period. This results in a  $^1\text{H}$ -coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum with resolved  $^{15}\text{N}$  spin doublet components. In this simple implementation of the experiment, both components of the doublet are in-phase. The cross-correlation term,  $\eta_{xy}$ , can then be determined directly from fitting the time dependence of the ratio of these signals to a mono-exponential decay function (Eq. 4.2.10). However, for large or partially unfolded proteins, overlap of peaks becomes a concern since  $^1\text{H}$ - $^{15}\text{N}$  scalar coupled spectra have twice as many peaks as decoupled spectra.

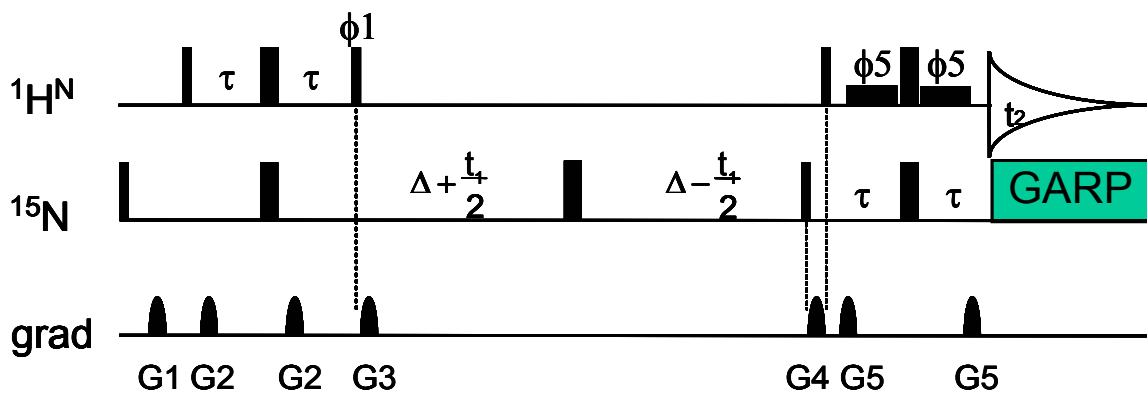
#### 4.2.2 IPAP Method for Simplification of Coupled Spectra

In 1998, Ottiger et al. introduced a method for the simplification of coupled spectra as an aid in the determination of J- and dipolar couplings from these spectra. This method, which is called “IPAP” for In-Phase Anti-Phase, involves acquiring two  $^1\text{H}$ - $^{15}\text{N}$  scalar coupled spectra, one spectrum in which the two peaks of the nitrogen doublet have the same phase and one in which the two peaks have opposite phase<sup>61</sup>. Experimentally, the reversal of the phase of the up-field component is brought about through the introduction of three pulses (shown in white in Fig. 4.2.2, we refer to these pulses as the “AP” element (for Anti-Phase) while their absence is referred to as the “IP” element (for In-Phase)). The addition of the processed in-phase and anti-phase spectra results in one spectrum (the “summation” spectrum) which only has

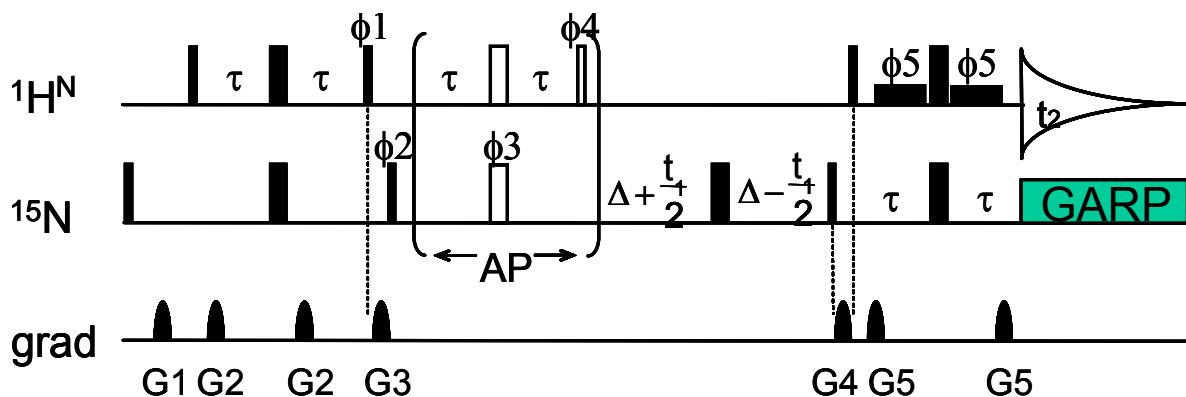


peaks corresponding to downfield components. Similarly, the subtraction of the processed anti-phase spectrum from the processed in-phase spectrum results in a simplified spectrum (the “difference” spectrum) which has only the up-field components. This addition and subtraction process is shown schematically in Fig. 4.2.3 and explained in detail below. The analysis of these simplified spectra eliminates problems due to signal overlap introduced by not suppressing the  $^1\text{H}$ - $^{15}\text{N}$  scalar coupling.

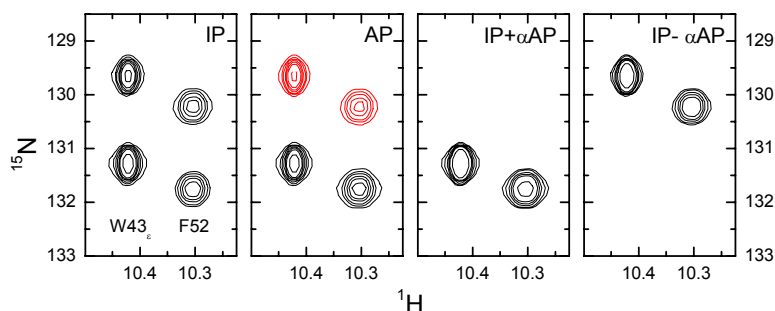
**Figure 4.2.1** Pulse sequence for the measurement of transverse  $^{15}\text{N}$  CSA/dipole-dipole cross-correlation rate from a coupled HSQC spectrum. IP (in-phase experiment) Narrow and wide pulses correspond to  $90^\circ$  and  $180^\circ$  flip angles respectively. The two low power pulses flanking the last nonselective  $180^\circ$  pulse on proton (as part of the water suppression element) are 1 ms long. The delay  $\tau$  is set to 2.5 ms ( $\sim 1/4J_{\text{NH}}$  for amides), and the relaxation period  $2\Delta$  varies from one experiment to another. The phases are  $\phi_1 = -y, y, \phi_3 = 4\{x\}, 4\{y\}, 4\{-x\}, 4\{-y\}, \phi_4 = 8\{x\}, 8\{-x\}, \phi_5 = -x, \phi_2 = 2\{x\}, 2\{-x\}$  and the receiver  $= x, -x, -x, x$ . For quadrature detection, phases  $\phi_2$  and  $\phi_3$  are incremented in the States-TPPI fashion. All other pulses are along  $x$ . Gradients were sine-shaped with the following strengths:  $G1=12 \text{ G cm}^{-1}$ ,  $G2=9 \text{ G cm}^{-1}$ ,  $G3=18 \text{ G cm}^{-1}$ ,  $G4=11 \text{ G cm}^{-1}$ ,  $G5=24 \text{ G cm}^{-1}$ . Their durations were 600  $\mu\text{s}$ , 600  $\mu\text{s}$ , 700  $\mu\text{s}$ , 600  $\mu\text{s}$  and 700  $\mu\text{s}$  respectively.



**Figure 4.2.2** Pulse scheme for the measurement of transverse  $^{15}\text{N}$  CSA/dipole-dipole cross-correlation rates. AP (anti-phase) experiment. The AP element of the pulse sequence (as discussed in text) is shown in parentheses. The phases are  $\phi_1 = -y, y, \phi_3 = 4\{x\}, 4\{y\}, 4\{-x\}, 4\{-y\}, \phi_4 = 8\{x\}, 8\{-x\}, \phi_5 = -x, \phi_2 = -y, -y, y, y$  and the receiver  $= x, -x, -x, x, -x, x, x, -x$ . Phases  $\phi_2$  and  $\phi_3$  and all gradients are as in 3.2.1



**Figure 4.2.3** Sections of spectra recorded using the sequence in Figure 4.2.1 for the IP (labeled top right-hand corner) and Figure 4.2.2 AP versions, and then a section of the spectra that are the result of the addition and subtraction (simplified spectra). The red and black contours represent positive and negative intensities, respectively. The scaling factor,  $\alpha$ , was optimized for cancellation of the residual intensities in the regions of the spectra where cancellation was expected by least-squares minimization. The value of  $\alpha$  in each experiment was determined from 49 non-overlapping doublets analyzed simultaneously. In the  $\eta_{xy}$  experiment which is presented in Figure 4.2.5, the values of  $\alpha$  were between 1.067 and 1.070.



In the IPAP method two spectra are recorded with the  $^{15}\text{N}$  doublet being in-phase (see above) and anti-phase (AP), and the two are added or subtracted to produce simplified spectra in which only one of the two components is retained while the other is eliminated. When the AP element is introduced into the pulse sequence (Fig.4.2.2), the corresponding signals can be written as  $-f\sigma_{\text{up}}$  and  $f\sigma_{\text{dn}}$ , where  $f$  represents signal attenuation due to the AP element. To compensate for these losses and to achieve full cancellation of the unwanted signals, an empirically determined scaling factor  $\alpha$  is applied to the AP spectrum prior to its addition to or subtraction from the IP spectrum. The ratio of the signals observed in the difference and sum spectra is then  $\sigma_{\text{diff}}/\sigma_{\text{sum}} = (\sigma_{\text{up}} + \alpha f \sigma_{\text{up}}) / (\sigma_{\text{dn}} + \alpha f \sigma_{\text{dn}}) = \sigma_{\text{up}}/\sigma_{\text{dn}}$ . This ratio of the

signals derived from the IPAP experiment is independent of the factors  $f$  and  $\alpha$ , which eliminates any possible bias by an arbitrary scaling factor, even if  $\alpha$  is not exactly set to  $1/f$ .

Both pulse sequences in Figures 4.2.1 and 4.2.2 make use of a constant-time spin evolution, in which signal evolution as a function of the incremented delay  $t_1$  is not directly modulated by spin relaxation<sup>110</sup>. The lineshapes here strongly depend on signal apodization in  $t_1$ . Therefore this sequence is subject to “wiggles” (base line oscillations) due to signal truncation if particular attention is not taken in selection of the window function applied in the  $t_1$  dimension. An extensive analysis indicates that of the standard set of window functions available within the XWINNMR package, squared *sinc* function provided the best results for GB3 spectra.

We applied the sequences in Figures 4.2.1 and 4.2.3 to measurement of the transverse  $^{15}\text{N}$  CSA/dipolar cross-correlation rates in the GB3 domain. Representative mono-exponential fits (Eq.4.2.10) for a few residues using both peak intensities and peak volumes are shown in Figure 4.2.4. The  $\eta_{xy}$  values determined here directly from the  $^1\text{H}$ -coupled spectrum (IP experiment) are in good agreement with those derived from the indirect A/B method<sup>43</sup> (Fig. 4.2.5). The comparison provided a scaling factor of 1.07 to correct the results of the A/B method. A comparison of the  $\eta$  values derived from the coupled spectrum (IP) with those from the simplified spectra from the IPAP experiment (Fig. 4.2.5) demonstrates that the IPAP scheme does not introduce any bias in the data. These comparisons indicate that the suggested approach is an accurate method for measuring the  $^{15}\text{N}$  CSA/dipolar CCRs.

Figure 4.2.4 Representative decay curves for the ratios of (a) Representative decay curves for the ratios of (a) peak volumes and (b) peak intensities from the IP experiment and (c) of peak intensities in the IPAP experiment. Shown are data for residues A34 (circles), T53 (triangles), and W43<sub>e</sub> (squares). The error bars are comparable to the size of the symbols. The corresponding  $\eta$  values for these residues are  $4.45 \pm 0.03 \text{ s}^{-1}$ ,  $3.45 \pm 0.02 \text{ s}^{-1}$ , and  $2.71 \pm 0.01 \text{ s}^{-1}$  derived from the ratios of peak volumes and  $4.50 \pm 0.01 \text{ s}^{-1}$ ,  $3.41 \pm 0.01 \text{ s}^{-1}$ , and  $2.71 \pm 0.01 \text{ s}^{-1}$  from peak intensities, and  $4.49 \pm 0.04 \text{ s}^{-1}$ ,  $3.41 \pm 0.01 \text{ s}^{-1}$ ,  $2.67 \pm 0.02 \text{ s}^{-1}$  from IPAP peak intensities.

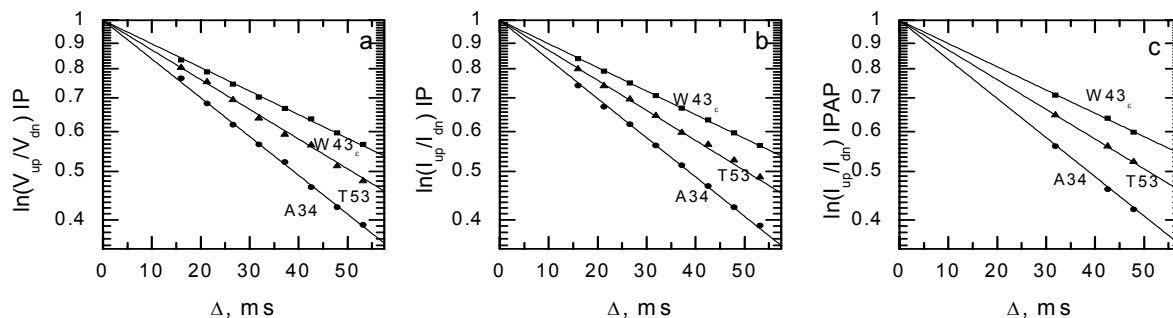
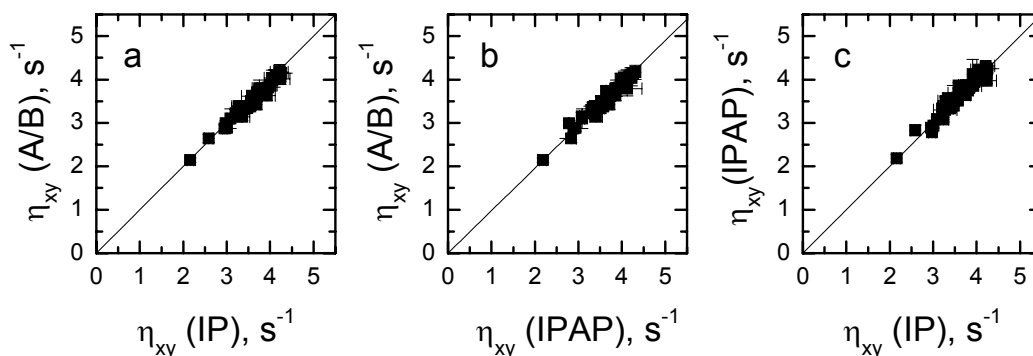


Figure 4.2.5 The agreement between the  $\eta_{xy}$  values measured here using (a) A/B experiment (Tjandra, J. Am. Chem. Soc., 1996) versus the IP method, (b) A/B method versus the IPAP scheme, and (c) IPAP versus the IP method. Only those spin systems (49 out of 56) that give isolated doublets were selected for the comparison with the IP data. The corresponding relaxation delays ( $\Delta$ ) for the measurements using the A/B method were set to 31.91, 42.55, 53.19, and 63.82 ms. The data points in panels a and b fall on a straight line with the slope less than one ( $0.934 \pm 0.020$  and  $0.930 \pm 0.018$ , respectively, correlation coefficient  $R = 0.96$ ), indicating a slight (7%) underestimation of the  $\eta$  values from the A/B method. This is a result of the difference in pulse sequences used in the experiments A and B and can be corrected by applying a uniform scaling factor of 1.07 to the  $\eta$  valued derived by the A/B method.



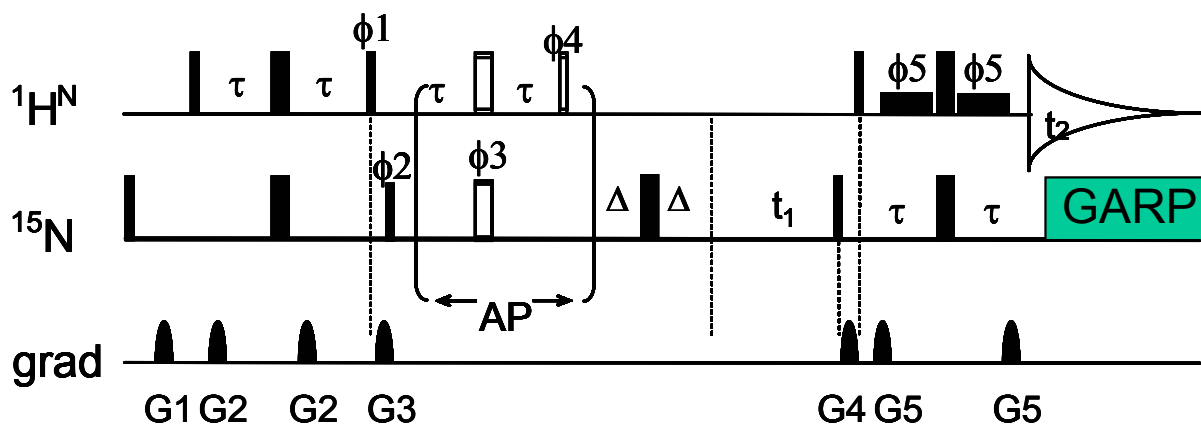
### 4.2.3 Real-time CCR delay experiment

As discussed above, the use of constant-time spin evolution has advantages and disadvantages for CCR measurement. Specifically, the constant time experimental scheme has the advantage that the signal linewidth in the  $^{15}\text{N}$  dimension is not directly related to  $^{15}\text{N}$  transverse relaxation, and, therefore, is less sensitive to line broadening in large proteins. However, the very small amount of signal decay as a function of  $t_1$  during the ct-evolution period can result in severe truncation artifacts in the F1 dimension, so the quality of the resulting spectra strongly depends on the applied window function. A mismatched apodization function could result in “wiggles” (base line oscillations) due to signal truncation in the  $^{15}\text{N}$ -dimension which can affect the amplitudes of  $^{15}\text{N}$  doublet components. Therefore, for applications to large proteins where relaxation broadening is a concern, the constant time scheme is the best choice. However, for small proteins and other molecules, where the  $^{15}\text{N}$  linewidths are not prohibitively large, a conventional (non-constant time)  $t_1$ -evolution period significantly reduces possible truncation artifacts.

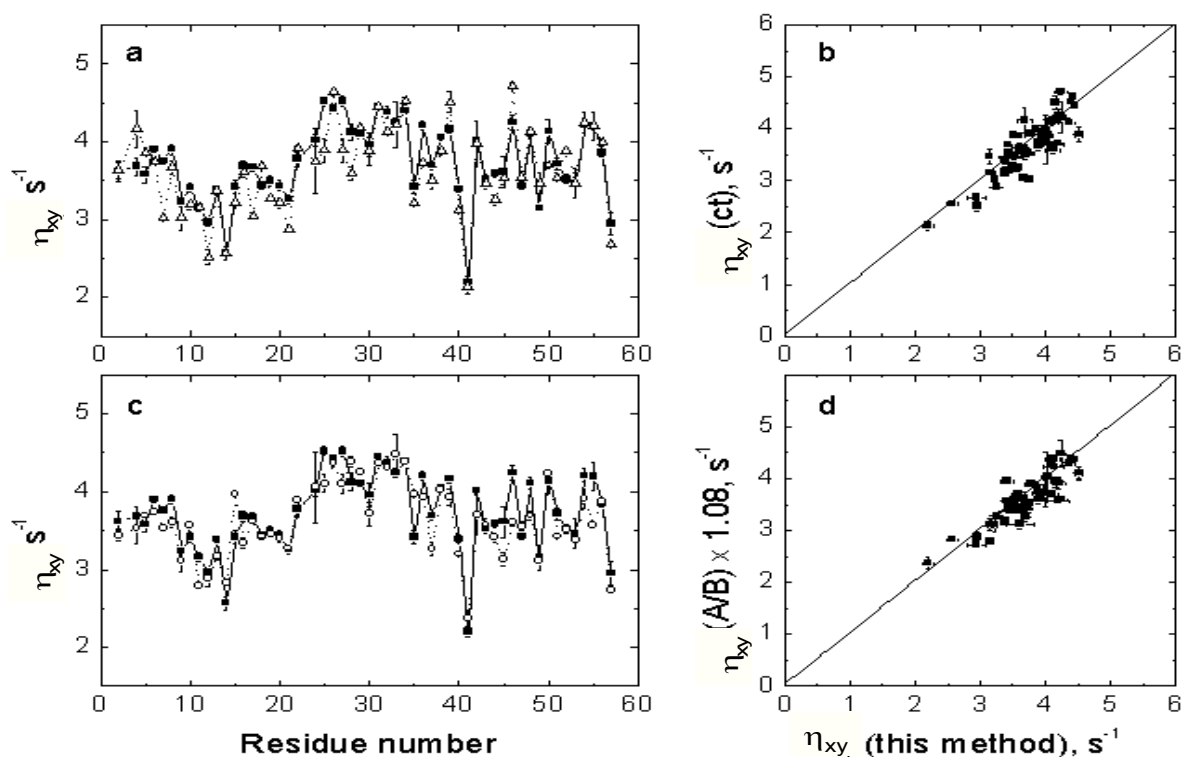
Figure 4.2.6 shows a pulse sequence for measurement of transverse  $^{15}\text{N}$  CSA/ $^1\text{H}$ - $^{15}\text{N}$  dipolar interference effects from  $^1\text{H}$ -coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC using the conventional  $t_1$ -evolution period. As in the constant-time method, the signal overlap problem in the coupled spectra is addressed by using the IPAP scheme to simplify the coupled HSQC spectra. Application of this technique to the B3 domain of protein G shows that this method also provides accurate measurements of the  $^{15}\text{N}$  CSA/dipolar cross-correlation rates (Figure 4.2.7). The CCRs from these conventional  $t_1$ -evolution

experiments are in agreement with the measurements using the IP and IPAP schemes and with corrected rates from the A/B experiment. Here the correction factor for the A/B experiment, obtained from a least-squares fit of the data to a linear model, is 1.08. Though in panel d of Figure 4.2.7 it looks by eye as though there is some systematic offset of the A/B data when this 1.08 correction factor is applied, this factor minimizes the fit of the A/B data to the real time CCR experiment data.

**Figure 4.2.6** Pulse sequence for the measurement of transverse  $^{15}\text{N}$  CSA/dipole-dipole cross-correlation rate from a coupled HSQC spectrum using conventional (not ct)  $t_1$  evolution. IP and AP experiments (the AP element is left out of the pulse sequence when running the IP experiment) are shown. All phases and gradients are as in Figs 4.2.1-4.2.2. The delay  $\tau$  is set to 2.5 ms ( $\sim 1/4J_{\text{NH}}$  for amides),  $\Delta$  is set to X, and  $t_1$  is varied.



**Figure 4.2.7** The agreement between  $\eta_{xy}$  values measured using the experimental scheme shown above and those determined (a,b) from the constant time experiment and (c,d) using the A/B method. Panels (a) and (c) present the comparison on a per residue basis: ct-data are shown as open triangles (a) and the A/B data as open circles (c), while the data obtained using conventional  $t_1$  evolution are shown as solid squares. The correlation coefficient in (b) is  $r=0.92$  and in (d) is  $r=0.93$ . A somewhat greater spread of the data points in (b) is due to truncation-related errors in the  $\eta_{xy}$  values from the ct-experiment. The data from the A/B method shown here were multiplied by a scaling factor 1.08, obtained from a linear least-squares fit of the A/B CCR rates to the real-time IPAP CCR experiment rates.



#### 4.2.4 Influence of Additional Relaxation Mechanisms

The expressions for the relaxation developed above assumed that the  $^{15}N$ - $^1H$  spin pair was isolated from all other spins and that all the relaxation was due to  $H_{DD}$  and  $H_{CSA}$  as defined in Chapter 2. Proteins, however, are spin dense systems, where

the  $^{15}\text{N}$  and  $^1\text{H}$  spins are coupled to many other protons within the molecule. If we represent the relaxation coefficients calculated above for the isolated case with subscript 0, and assume that the additional relaxation mechanisms act independently on the nitrogen and proton nuclei (i.e. there are no interference effects (CCRs) produced by these additional mechanisms),

$$R_2^{av} = R_{2,0}^{av} + \frac{1}{2} R_{2,N}' + \frac{1}{2} R_{2N_yH_z}' = R_{2,0}^{av} + \frac{1}{2} R_{2,N}' + \frac{1}{2} R_{2,N}' + \frac{1}{2} R_{1,H}' = R_{2,0}^{av} + R_{2,N}' + \frac{1}{2} R_{1,H}' \quad (4.2.12)$$

$$R_2^{\frac{1}{2}dif} = R_{2,0}^{\frac{1}{2}dif} + \frac{1}{2} R_{2,N}' - \frac{1}{2} R_{2N_yH_z}' = R_{2,0}^{\frac{1}{2}dif} + \frac{1}{2} R_{2,N}' - \frac{1}{2} R_{2,N}' - \frac{1}{2} R_{1,H}' = R_{2,0}^{\frac{1}{2}dif} - \frac{1}{2} R_{1,H}' \quad (4.2.13)$$

where the primes indicate additional relaxation caused by additional spins. Eqs. 4.2.7 and 4.2.8 taking into account the additional relaxation contributions can then be expressed as:

$$\frac{d}{dt} \langle N_y^{1H(up)} \rangle = -i \frac{J_{NH}}{2} \langle N_y^{1H(up)} \rangle - (R_{2,0}^{av} + R_{2,N}' + \frac{1}{2} R_{1,H}' + \eta_{xy}) \langle N_y^{1H(up)} \rangle - (R_{2,0}^{\frac{1}{2}dif} - \frac{1}{2} R_{1,H}') \langle N_y^{1H(dn)} \rangle \quad (4.2.14)$$

$$\frac{d}{dt} \langle N_y^{1H(dn)} \rangle = i \frac{J_{NH}}{2} \langle N_y^{1H(dn)} \rangle - (R_{2,0}^{av} + R_{2,N}' + \frac{1}{2} R_{1,H}' - \eta_{xy}) \langle N_y^{1H(dn)} \rangle - (R_{2,0}^{\frac{1}{2}dif} - \frac{1}{2} R_{1,H}') \langle N_y^{1H(up)} \rangle \quad (4.2.15)$$

Now the relaxation rates of the individual components of the doublet are given by:

$$R_2^{1H(up)} = R_2^{av} + \eta_{xy} + R_{2,N}' + \frac{1}{2} R_{1,H}'^{(up)} \quad \text{and} \quad R_2^{1H(dn)} = R_2^{av} - \eta_{xy} + R_{2,N}' + \frac{1}{2} R_{1,H}'^{(dn)} \quad (4.2.16)$$

where  $\frac{1}{2} R_{1,H}'^{(up)}$  is not strictly equal to  $\frac{1}{2} R_{1,H}'^{(dn)}$ . Kay et. al<sup>111</sup> consider the term  $\frac{1}{2} R_{1,H}'$ , as the half exchange-rate between the two doublet components due to spontaneous spin



flips of the proton (proton  $R_1$  relaxation). This is clearly in the slow-on-the-NMR-timescale exchange regime, since the splitting between the two components of the doublet  $\frac{J_{NH}}{2} \sim 94\text{Hz}$ , is much greater than  $\frac{1}{2}R'_{1,H} \sim 4 - 5\text{Hz}$ . Therefore as long as the rate of exchange  $k_{ex}$ , for transition for the proton spin,  $H_z$  from spin-up  $\rightarrow$  spin-down is equal to the rate for spin-down  $\rightarrow$  spin-up, the decay of the ratio of the intensity of the upfield peak to the intensity of the downfield peak is still given by Eq. 4.2.11. Thus the time evolution of the ratio of the two components is insensitive to proton  $R_1$  relaxation.

### **4.3 Spin State Selection for Measurement of CSA/dipolar CCRs**

A mono-exponential fit of the time evolution of the ratio of the intensities of the up- and down-field components of the nitrogen doublet using either the constant-time evolution or the conventional  $t_1$  evolution schemes presented above, is the most straightforward way to obtain  $^{15}\text{N}$  CSA/dipolar CCRs. In large proteins, where spectral overlap is an issue, the IPAP method can be used to simplify the spectra. However, the addition and subtraction of the two spectra obtained from the in-phase and anti-phase experiments can be problematic in the instance of severe peak overlap. As discussed above, a small correction factor,  $f_0$ , is used to compensate for any differences between the IP and AP experiments prior to their linear combination. It can be shown that errors in restored-peak intensities can arise from differences in relaxation properties of amides, resulting in a difference between the overall AP-correction factor  $f_0$  and the signal-specific correction factors ( $f_1$  and  $f_2$ ) for two overlapping signals (here designated  $S_1$  and  $S_2$ ). The relative error in the ratio of peak

intensities of the up- and downfield components of  $S_1$  can be approximated as  $\delta(S_{1up}/S_{1dn})_{IPAP} \approx \frac{1}{2} \chi(S_{1up}/S_{1dn})(f_2 - f_0)$ , where  $\chi$  is the degree of overlap (i.e. the error in  $S_{1up}$  due to the overlap is  $\delta S_{1up} = \chi S_{2dn}$ ), and we assumed that  $f_1, f_2, f_0 \sim 1$ . Considering that the difference between the overall and the signal-specific correction factors in GB3 is at most  $\pm 5\%$ , the error introduced in the IPAP approach is considerably smaller than that introduced by untreated overlap (where  $\delta(S_{1up}/S_{1dn})_{IP} = \chi S_{1up}/S_{1dn}$ ). It can be concluded that CCR measurements via coupled HSQC experiments introduce relative errors in the signal ratios of the order of the degree of overlap for in-phase coupled spectra analyzed alone, and about or less than 2.5% of the degree of overlap in IPAP experiments. Because of the different relaxation rates of  $S_{2dn}$  and  $S_{1up}$ , their ratio will depend on the CCR delay  $2\Delta$ , which in turn could affect the measured values of  $\eta$ .

One way to resolve this is the use of a spin-state selection method, to select the individual components of the  $^{15}\text{N}$  doublet prior to the relaxation period. Though less “direct” than the coupled HSQC method for CCR measurement, we have shown that this spin-state selection method is an improvement over the IPAP method for severely overlapping signals<sup>112</sup>. This new direct method for measuring transverse cross-correlation rates selects coherences (single-transition operators) corresponding to a given component of the  $^{15}\text{N}$  doublet at the beginning of the CCR-decay period. The separate relaxation rates of the two components,  $R_{2\pm\eta}$ , can thus be determined directly. Furthermore, the observed spectra are “simplified” (compared to a  $^1\text{H}$ -coupled HSQC) as the number of signals is the same as in the decoupled spectrum.

In order to generate the single-transition  $^{15}\text{N}$ -operators we used the spin-state-selective element ( $\text{S}^3\text{E}$ ) of Sørensen et al.<sup>62</sup>. The  $\text{S}^3\text{E}$  selects either the  $\sigma_{\text{up}}$  or the  $\sigma_{\text{dn}}$  component at the beginning of the CCR delay  $2\Delta$ , depending on the phase cycle chosen. Alternative spin-state selection modules exist in the literature<sup>113 114 115</sup>, we have chosen the  $\text{S}^3\text{E}$  building block because it can yield data with similar sensitivity to that of the IPAP sequence, when the scans for each of the selection phase cycles are stored separately and subsequently processed in linear combinations.

The application of the  $\text{S}^3\text{E}$  filter-CCR period- $^{15}\text{N}$ -evolution-without-decoupling formula also has an advantage over other pulse sequence schemes using selection elements, in that selection of the desired component can be monitored and clean selection ensured by adjusting critical parameters of the sequence: the  $\text{S}^3\text{E}$  delay  $2\delta$ , and the  $^{15}\text{N}$  pulses (especially the  $180^\circ$  pulse in the middle of the selection filter). For example, a delay  $2\delta$  in the selection filter different from the optimal value for this parameter,  $2\delta_{\text{opt}}=1/(4J)$ , leads to a decrease of the selected component by a factor of  $\cos(2\pi J(\delta-\delta_{\text{opt}}))$  and introduces an artifact signal in the spectra at the position of the unwanted component, with the intensity proportional to  $-\sin(2\pi J(\delta-\delta_{\text{opt}}))$ . The sign of the artifact is the same for both selected components and varies from positive to negative as a function of  $\delta - 1/(4J)$ . Another critical factor for proper coherence selection is pulse calibration, especially the calibration of  $^{15}\text{N}$  pulses. The  $\text{S}^3\text{E}$  selection module uses composite  $180^\circ$  pulses ( $90^\circ_y$ - $180^\circ_x$ - $90^\circ_y$ ). A miscalibrated  $180^\circ_x$   $^{15}\text{N}$  pulse in the middle of the composite pulse results in a decrease in the intensity of the selected component by a factor  $\cos(\alpha)$ , where  $\alpha$  is the deviation of the corresponding flip angle from  $180^\circ$ . In addition, spectral artifacts appear at the

position of the unwanted component, with intensities proportional to  $\sin(\alpha)$  if the upfield component is selected and  $-\sin(\alpha)$  in the case of the downfield component. Consequently, artifacts caused by the imperfection of the  $180^\circ$ -pulse in the middle of the composite pulse can be identified because they change sign depending on whether the pulse is longer or shorter than its ideal value. Within the approximation that the relaxation matrix is diagonal in the representation of single-transition operators, the presence of an unwanted signal belonging to the complementary component does not affect the measurement of the CCR, errors are only introduced when this additional signal overlaps with that of another amide.

The  $S^3E$  CCR experiment was tested on GB3 at 500 and 600 MHz. The results are in good agreement with those obtained from the IP/IPAP coupled HSQC experiments presented above (Fig 4.3.2). There is expected disagreement for overlapping residues; seven residues in GB3 show overlap only in the coupled spectra: K4,V5,A23,K24,A31,D47,A48 at 600 MHz and Q2,K4,V5,A23,D47,A48,V54 at 500 MHz (in green in Figure 4.3.2 a,b); the peaks were considered overlapping if their centers were separated by less than 0.6 ppm in  $^{15}N$  and 0.06 ppm in the  $^1H$  dimension. Of these residues, A23 was most affected by the overlap: neither IP nor IPAP data (600 MHz) could be fitted well to an exponential decay, whereas the  $S^3E$  data fit well. The results for the other overlapping residues follow the expected trend, with IPAP data in better agreement with the  $S^3E$  results than the results of the IP-only method.

**Figure 4.3.1** S<sup>3</sup>E-selective pulse sequence for measuring transverse <sup>15</sup>N CSA/dipolar cross-correlation rates using both conventional (A) and constant time t<sub>1</sub> evolution (B). Here the open bars represent composite 90<sub>y</sub>-180<sub>x</sub>-90<sub>y</sub> pulses. The delays are: τ=2.6 ms, δ=1.22ms, the duration of the CCR-relaxation delay 2Δ is set to either zero or multiples of 1/J. The relative intensities of the gradients are G1:G2:G3=1:2.3:1.4. Two experiments are run with different phase cycles. The two spectra are then added (subtracted) to yield the downfield (upfield) components. The phase cycling for the first experiment is φ<sub>1</sub>=x,-x, φ<sub>2</sub>=4(45°),4(225°), φ<sub>3</sub>=2(x),2(y); φ<sub>4</sub>=2(x),2(y), φ<sub>5</sub>=8(x),8(-x), φ<sub>rec</sub>=x,2(-x),x,-x,2(x),2(-x),2(x),-x,x,2(-x),x while for the second experiment φ<sub>2</sub>=2(2(45°),2(225°)) and φ<sub>4</sub>=2(x),2(-y).

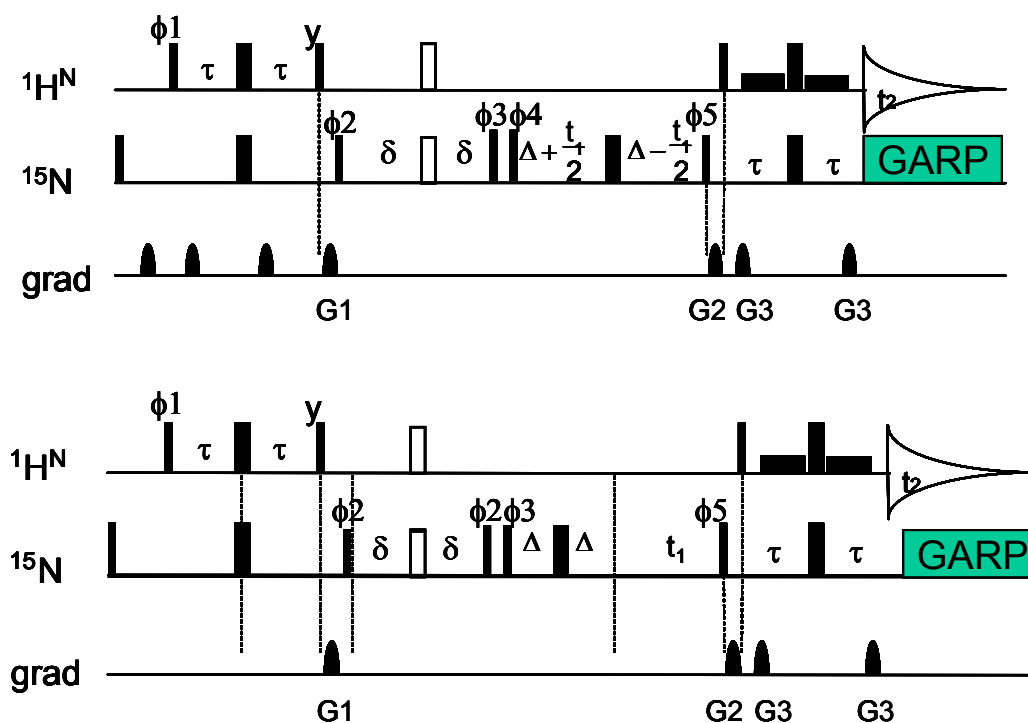
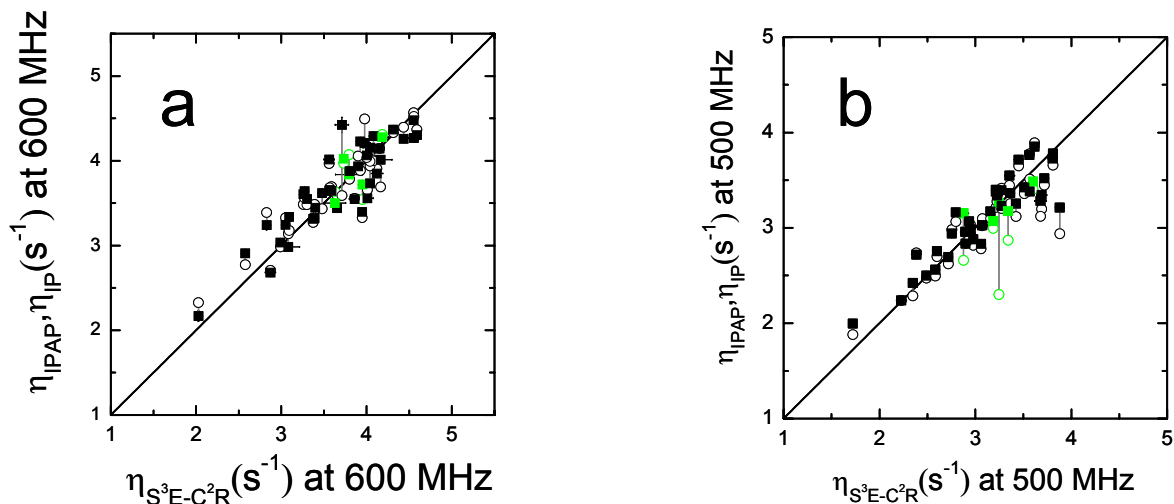


Figure 4.3.2 Correlation between  $\eta_{xy}$  measured at 14.1 T using the IP method (open circles) and using the IPAP method (filled squares) with  $\eta_{xy}$  measured using the S3E method (x-axis). The IP and IPAP-derived values for each residue are connected by vertical lines. The correlation coefficient is 0.88 (S<sup>3</sup>E vs. IPAP) and 0.91 (S<sup>3</sup>E vs. IP) (for non-overlapping residues). The error bars reflect stochastic errors, calculated on the basis of the S/N ratio of the spectra. The data points for residues affected by signal overlap in the coupled spectra are colored green.



#### 4.4 Longitudinal CSA/dipolar CCR

An approximation was made in writing Eqs. 4.2.1, 4.2.7 and 4.2.8, namely that the  $^{15}\text{N}$ - $^1\text{H}$  spin pair was isolated from all other spins and that all the relaxation was accounted for in the Hamiltonians,  $H_{\text{DD}}$  and  $H_{\text{CSA}}$ . In sec 4.2.4, this approximation was lifted, and the relaxation equations for transverse relaxation of the up- and down-field components of the nitrogen doublet in the presence of other relaxation mechanisms were given in Eqs. 4.2.13-4.2.15. For longitudinal CSA/dipolar cross-correlated relaxation, the analogous relaxation equations for the isolated spin system are given by:

$$\frac{\partial}{\partial t} \begin{pmatrix} \langle N_z(t) \rangle \\ \langle 2N_z H_z(t) \rangle \end{pmatrix} = \begin{pmatrix} R_{1,N} & \eta_{z,N} \\ \eta_{z,N} & R_{1,2HN} \end{pmatrix} \cdot \begin{pmatrix} \langle N_z(t) \rangle \\ \langle 2N_z H_z(t) \rangle \end{pmatrix}, \quad (4.1.1)$$

where  $R_{1,N}$ , and  $R_{1,2HN}$  are the longitudinal auto-relaxation rates of nitrogen and two-spin order longitudinal magnetization respectively,  $\eta_{z,N}$  is the longitudinal

CSA/dipolar CCR of nitrogen. Longitudinal cross-correlated relaxation is slightly more complicated than transverse cross-correlated relaxation because the  $N_z$  and  $2N_zH_z$  operators commute with the scalar coupling Hamiltonian. However, if these components of magnetization can be averaged by experimental methods<sup>101</sup>, the relaxation of the components will be given by:

$$R_1^{1H(up)} = \left( \frac{R_{1,N} + R_{1,2HN}}{2} \right) + \eta_z \text{ and } R_1^{1H(dn)} = \left( \frac{R_{1,N} + R_{1,2HN}}{2} \right) - \eta_z \quad (4.2.6)$$

$$\frac{\sigma^{up}}{\sigma^{dn}} = \frac{e^{-(\bar{R}_1 + \eta_z)t}}{e^{-(\bar{R}_1 - \eta_z)t}} = e^{-2\eta_z t},$$

for the longitudinal CCR experiment, with  $\bar{R}_1$  as  $\left( \frac{R_{1,N} + R_{1,2HN}}{2} \right)$ .

However, there is an additional consideration when measuring the longitudinal CCR since proton spin diffusion may contribute to measured values of longitudinal cross-correlation relaxation<sup>101,116</sup>. To examine the magnitude of this effect on longitudinal CCR's in GB3 we measured  $\eta_z$  in both a protonated and deuterated GB3 sample and obtained results that were identical within the estimated errors of the measurement. Though the  $\eta_z$  values in the deuterated sample are systematically lower than those in the protonated sample, this difference is within the experimental error of the measurement. We therefore conclude that the effect of spin diffusion on  $\eta_z$  can probably be neglected in this system. Figure 4.4.1 shows the pulse sequence for  $\eta_z$  measurement and Figure 4.4.2 shows the measured  $\eta_z$  values in protonated and deuterated samples of GB3. The possibility that the small difference between these measurements is systematic is still being explored.





## **4.5 Conclusions**

Direct methods for measuring the  $^{15}\text{N}$  CSA/dipolar cross-correlation rates in proteins are in good agreement with indirect methods (Figures 4.2.5b and 4.2.7d), but do not need to be corrected for scaling factors introduced by small differences in experimental conditions or pulses between the two experiments required for indirect measurement of the build-up of  $2N_yH_z$  coherence from  $N_y$  coherence. We have shown here that  $^{15}\text{N}$  CSA/dipolar cross-correlation rates can be measured directly from the relative amplitudes of the up- and down-field  $^{15}\text{N}$  signals in a  $^1\text{H}$ -coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum. This type of measurement of CCRs from the ratio of intensities of doublet components is quite general for measuring cross-correlations for other interactions. Examples include measurements of cross-correlation rates between  $\text{H}^{\text{N}}$  CSA and  $\text{H}^{\text{N}}$ - $^{15}\text{N}$  dipolar coupling and between  $^{13}\text{CO}$  CSA and  $^{13}\text{CO}$ - $^{13}\text{C}\alpha$  dipolar interactions<sup>73</sup>. The obvious advantage of this approach is that both signals are observed in the same spectrum and, therefore, no ambiguity associated with correction factors is involved. The application of this method to biological macromolecules, however, is complicated by signal overlap in the coupled 2D spectra, which may be particularly severe in the case of  $\text{H}$ -coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for macromolecules greater than 10 kDa.

We have shown that the IPAP scheme simplifies coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra without causing deviations in cross-correlation rates. Both CCRs measured using the IP only sequence and the IPAP scheme agree with the indirect A/B method when the A/B method is scaled by a correction factor (Fig 4.2.5b).

$S^3E$  spin-state selection of the individual (up- or down-field) component of the nitrogen doublet before the relaxation delay also alleviates problems due to spectral overlap. This correction is, in principle, even better than that of the IPAP scheme which can introduce very small errors in restored-peak intensities due to site-specific differences in relaxation properties of individual amides. CCRs measured using the sequence containing the  $S^3E$  spin-state selection element agree with IPAP, IP, and scaled A/B method CCRs.

In subsequent chapters I will demonstrate the utility of  $^{15}\text{N}\{-^1\text{H}\}$  CSA/dipolar CCRs to identifying conformational exchange motions, for determining the overall diffusion properties of a molecule, and for the determination of site-specific  $^{15}\text{N}$  CSAs in proteins. All of these applications rely on the accuracy and precision of the CCR rates, therefore the techniques presented here are valuable tools for improving our ability to probe both protein dynamics and protein chemistry.

## Chapter 5: Overall Rotational Diffusion Tensor of GB3

### 5.1 Motivation

Overall rotational diffusion is generally a larger cause of spin relaxation than fast, local motion of individual bonds or collective motions of groups of atoms<sup>34</sup>. If it was not for the fact that proton relaxation in the nuclear Overhauser effect can be approximated as arising from the Brownian motion of a rigid molecule, current methods of NMR structure determination of proteins would not be possible. Since the overall tumbling has a much larger effect on nuclear spin relaxation rates than the motion of the NH bonds, the overall tumbling must be correctly determined before any accurate picture of local NH bond motion can be deconvolved from NMR spin-relaxation rates. The dependence of spin relaxation rates on the angle between the dipole-dipole interaction and the symmetry axis for cylindrically symmetric rotational diffusion (also called axially symmetric or symmetric-top rotational diffusion) was first worked out by Woessner in 1962<sup>70</sup>. The experimental diffusion tensor of a molecule in solution can be determined by fitting experimental NMR spin relaxation rates to these theoretical expressions using a  $\chi^2$  minimization in the relevant parameter space. Whether a particular molecule is best described by an isotropic, cylindrically symmetric, or fully anisotropic rotational diffusion tensor can be evaluated using statistical criteria<sup>117</sup>.

However, this  $\chi^2$  minimization in the instance of cylindrically symmetric (4 parameter space) or fully anisotropic (6 parameter space) models of overall rotational diffusion requires considerably more computational effort than in the case of isotropic

overall diffusion (1 parameter space), and requires knowledge of the three-dimensional structure of the molecule. Early studies of protein dynamics using NMR spin relaxation rates, therefore, frequently relied on the assumption that the molecular rotational diffusion could be approximated as isotropic. Such a study of the local dynamics of NH bonds in the B1 domain of protein G (called GB1, sequence homology to GB3=96%, sequence identity to GB3=89%) which is structurally very similar to GB3 found conformational exchange motions throughout the  $\alpha$ -helix of the molecule<sup>118</sup>. These motions were attributed to “breathing” motions of the helix with respect to the  $\beta$ -sheet. Since the time of this study, several methodological studies of techniques for analyses of local motion from NMR relaxation data have pointed out that microdynamic parameters derived from these data could be in error if the rotational anisotropy is not correctly taken into account (as pointed out in e.g.<sup>119</sup>). Specifically, it has been shown<sup>44,63,64</sup>) that an analysis of <sup>15</sup>N relaxation data which does not include significant rotational anisotropy could result in spurious conformational exchange motions. Analysis of a representative set of 878 protein structures suggests that about 70% of monomeric proteins have  $1.2 < D_{\parallel}/D_{\perp} < 2$  (Geraghty et al, unpublished), which indicates that anisotropic rotational diffusion is quite general for proteins.

The shapes of both the GB3 and the GB1 domains display intermediate anisotropy; the inertia tensor of GB3 was calculated using the coordinates of the heavy atoms from the protein databank crystal structure (1IDG.pdb). The normalized values of the principal components of the inertia tensor are 1.80: 1.79: 1.00. It is therefore possible that the overall rotational diffusion of GB3 is significantly

anisotropic and that the assumption of isotropic rotational diffusion could result in overestimation of conformational exchange motions. Given this, it was of interest to determine the experimental diffusion tensor of GB3. This chapter describes the method and results of that determination, while the subsequent chapter discusses other experiments we conducted to identify and/or exclude conformational exchange motions in GB3.

## **5.2 Method for Derivation of the Rotational Diffusion Tensor of A Molecule from NMR Relaxation Data**

### **5.2.1 Method for Derivation of Parameters that Describe Rotational Diffusion from $^{15}\text{N}$ Relaxation Data.**

Anisotropic rotational diffusion of a molecule means that the rotation/reorientation of the molecule about some direction in space is faster than about other directions. This means that different  $^{15}\text{N}$  nuclei “feel” different overall correlation times, and have different spin-relaxation rates depending on the orientation of their dipolar interaction with  $^1\text{H}$  (their N-H bond vector) with respect to the principal axis frame of the rotational diffusion tensor. Given the structure of the molecule and experimental relaxation data ( $^{15}\text{N}$   $R_1$  and  $R_2$  and  $^{15}\text{N}\{^1\text{H}\}$  NOE), the overall diffusion tensor can be calculated by minimization of the target function:

$$\chi^2 = \sum_{i=1}^{N_r} \left( \frac{\rho_i^{\text{exp}} - \rho_i^{\text{calc}}}{\sigma_i} \right)^2 \quad (5.2.1)$$

where  $N_r$  is the total number of NH bond vectors in the analysis (usually the number of residues in the protein), and the parameter  $\rho^{\text{exp}}$  is the ratio of the reduced  $^{15}\text{N}$  transverse and longitudinal reduced relaxation rates (Eqs. 2.5.19-2.5.20):

$$\rho^{\text{exp}} = \left( \frac{2R_2'}{R_1'} - 1 \right)^{-1} = \frac{3}{4} \frac{J(\omega_N)}{J(0)}. \quad (5.2.2)$$

$\sigma_i$  denotes the experimental error in  $\rho_i$  for NH vector  $i$ .  $\rho^{\text{calc}}$  is calculated using expressions for the theoretical dependence of this ratio on the overall correlation time, the principal values of the overall rotational diffusion tensor, and the angles between the axes of the diffusion tensor and the NH bond vectors. The ratio,  $R_2'/R_1'$  (where  $R_1'$  and  $R_2'$  are defined in Eqs. 2.5.19-20), is used instead of the individual values of these parameters since this ratio is approximately independent of site-specific variations in  $r_{\text{NH}}$  and the  $^{15}\text{N}$  CSA. Furthermore, the  $R_2'/R_1'$  ratio is less sensitive to internal molecular dynamics<sup>120</sup>.

The theoretical expressions for this ratio in the approximation of no local motion (i.e.  $S^2 = 1$  so that all terms proportional to  $(1-S^2)$  can be ignored) and no chemical exchange, in the instances of isotropic and cylindrically symmetric rotational diffusion are given by<sup>121</sup>:

$$\text{Isotropic rotational diffusion: } \rho^{\text{calc}} = \frac{3}{4} (1 + \omega_N^2 \tau_c^2) \quad (5.2.3)$$

Cylindrical symmetry:

$$\rho^{\text{calc}} = \frac{3/4}{1 + (\omega_N \tau_1)^2} \left( 1 + \frac{(\omega_N \tau_1)^2}{(\omega_N \tau_1)^2 + (1 + \frac{1}{6} \varepsilon)^2} \frac{\varepsilon \sin^2 \theta}{3 + 2\varepsilon + \left[ 1 + \frac{1}{3} \varepsilon (2 - 3 \sin^2 \theta) \right]^2} x \right. \\ \left. \left[ 4 + 3\varepsilon + \frac{2}{9} \varepsilon^2 - \varepsilon \sin^2 \theta \left( 1 + \frac{4 + \frac{11}{3} \varepsilon + \frac{19}{18} \varepsilon^2 + \frac{5}{4} \varepsilon^3}{(\omega_N \tau_1)^2 + (1 + \frac{2}{3} \varepsilon)^2} \right) \right] \right) \quad (5.2.4)$$

where  $\theta$  is the angle between the NH vector and the axis of the diffusion tensor,  $\varepsilon = D_{\parallel} / D_{\perp} - 1$ ,  $\tau_1 = 6D_{\perp}$ , with  $D_{\parallel}$  and  $D_{\perp}$  representing the principal values of the diffusion tensor ( $D_{\perp} < D_{\parallel}$ ), and  $\omega_N$  the  $^{15}\text{N}$  Larmor frequency. In the case of axially symmetric rotational diffusion, the overall correlation time is given by

$$\tau_c = 1/2\text{tr}(D) = 1/2(D_{\parallel} + 2D_{\perp}). \quad (5.2.5)$$

Equation 5.2.4 can be calculated from equation 5.2.2 and  $J(\omega_N)$  and  $J(0)$  calculated from the expression for the spectral density function (in the case of no local motion) with an axially symmetric diffusion tensor, Eq. 3.4.14. The expression in the fully anisotropic model of rotational diffusion is more complicated. It can be worked out (but is difficult to write) using equation 5.2.2 and the expression for the spectral density function (in the case of no local motion) for a fully anisotropic tensor (Eq. 3.4.17).

In the most general case of a completely anisotropic diffusion tensor, six parameters have to be determined: the principal values of the tensor ( $D_x$ ,  $D_y$ ,  $D_z$ ) and the three Euler angles ( $\Phi$ ,  $\Theta$ ,  $\Psi$ ) that define the orientation of the principal axes frame of the tensor with respect to the molecular frame. In the axially symmetric case, the number of parameters is reduced to four:  $D_{\parallel}$ ,  $D_{\perp}$ , and the two Euler angles ( $\Phi$ ,  $\Theta$ ). The search of the parameter space for the solutions which minimize Equation 5.2.1 was performed using the computer program ROTDIF (developed in the lab) which uses a Levenberg-Marquardt minimization algorithm which has been shown to be much more efficient (and potentially more accurate) than grid-search methods<sup>120</sup>.

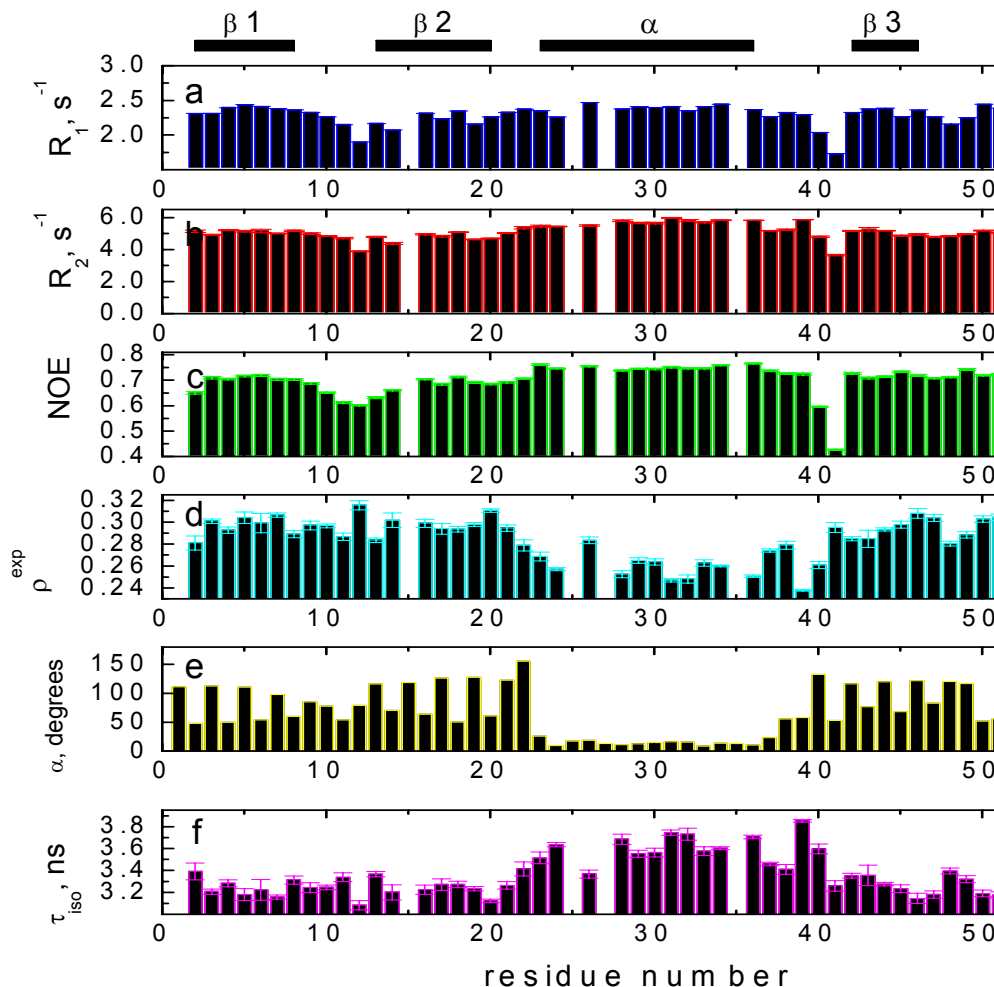
### **5.3 The Rotational Diffusion Tensor of GB3**

### 5.3.1 Relaxation Data

The Relaxation data used in this analysis consisted of the rates of  $^{15}\text{N}$  longitudinal ( $R_1$ ) and transverse ( $R_2$ ) relaxation and the rate of  $^{15}\text{N}$  -  $^1\text{H}$  cross-relaxation measured via the steady-state  $^{15}\text{N}$  {  $^1\text{H}$  } nuclear Overhauser effect (NOE) measured at 14.1 T and 24° C. The experiments were performed using the methods for relaxation rate measurement described in Chapter 3. 55 resolved backbone amide cross peaks were observed in the 2D spectra. Though they could be sufficiently resolved for assignment, residues Glu15 and Asn35 and Thr25 and Glu27 are not included in the analysis, as their signal intensities could be affected by spectral overlap. Gln2 (which is a mutation in our GB3 protein sample) was excluded from anisotropic analyses because it is not present in the available protein coordinates. The relaxation rates are shown in Fig. 5.3.1a-c.



**Figure 5.3.1. Relaxation Rates in GB3 at 600 MHz (14.1 Tesla).** (a)-(c) Amide  $^{15}\text{N}$  relaxation rates at 14.1 Tesla, (a)  $R_1$ , (b)  $R_2$ , and (c)  $^{15}\text{N}\{^1\text{H}\}$  NOE versus residue number for the B3 domain of protein G. The error bars represent standard errors in the experimental parameters. (d)  $\rho^{\text{exp}}$  values (Eq. 5.2.2) versus residue number, (e) the polar angle,  $\alpha$ , between each NH vector and the z-axis of the molecular frame, (f)  $\tau_{\text{iso}}$  calculated using Eq. 5.3.1 versus residue number. The horizontal bars on the top indicate the positions of the secondary structure elements in the protein sequence.



The  $R_1$ ,  $R_2$  and NOE data (Fig.5.3.1a-c) show simultaneous decrease in the  $\beta 1/\beta 2$  and  $\alpha/\beta 3$  loops, indicating that these are flexible regions in the protein. There is no such decrease in any of the relaxation parameters in the  $\beta 2/\alpha$  loop, however, and in the loop connecting strands  $\beta 3$  and  $\beta 4$  there is a decrease in  $R_1$  and  $R_2$  but not in the NOE. Noticeably elevated  $R_2$  values are observed for the entire  $\alpha$ -helix (Fig.

2b). The NOEs are also somewhat higher here than in the rest of the backbone, while the  $R_1$  values are at about the same level as in the other elements of the secondary structure. An elevation in  $R_2$  as observed in the  $\alpha$ -helix could be indicative of conformational exchange on the microsecond-millisecond timescale. However, given the structure of GB3, it is difficult to imagine a physical model that would account for every residue in the helix (including those not facing the  $\beta$ -sheet) being involved in motion on the  $\mu$ s-ms timescale. The orientation dependence of the transverse relaxation rate could account for the elevation in  $R_2$  for residues in the  $\alpha$ -helix if the helix axis of the GB3 domain is aligned parallel to the longitudinal axis of a prolate rotational diffusion tensor. This orientation would align the NH vectors in the  $\alpha$ -helix along the axis of fast overall rotation – as the result, they would experience slower rates of overall tumbling (hence higher  $R_2$ s) compared to the rest of the protein.

Figure 5.3.1d shows the experimental values of the parameter,  $\rho^{\text{exp}}$ , (Eq. 5.2.2) calculated from the relaxation rates. And Figure 5.3.1f shows the values of  $\tau_{\text{iso}}$  calculated from  $\rho_{\text{exp}}$  using Eq. 5.2.3 which assumes isotropic overall rotation:

$$\tau_{\text{iso}} = \frac{\sqrt{\frac{4}{3}\rho^{\text{exp}} - 1}}{\omega_N} \quad (5.3.1)$$

There is systematic variation (mean 3.36 ns, standard deviation 0.18 ns, difference between max and min  $\tau_{\text{iso}}$  of 0.77 ns) in these calculated values of  $\tau_{\text{iso}}$  from residue-to-residue, inconsistent with an isotropic model of overall diffusion, where all groups should experience the same overall correlation time. Figure 5.3.1e shows the values

of the polar angle,  $\alpha$ , of each NH vector in the crystal structure of GB3 (1IGD.pdb) with respect to the z-axis of the molecular frame. There is some correlation between these angles and the values of  $\rho^{\text{exp}}$  and  $\tau_{\text{iso}}$ , indicating that the protein does not tumble isotropically.

The residues with lower than average values of  $R_1$ ,  $R_2$ , and the NOE were excluded from the set of residues used to derive the anisotropic overall rotational diffusion tensors. As mentioned above, the low values of these rates indicate that the NH bonds in these residues are undergoing motion on the ps-ns timescale. Due to this motion, the orientation of the NH vectors for these residues in the crystal structure “snapshot” might not be representative of the time-averaged orientation on the time scale of the overall rotation (which is on the order of ns). Since this time-averaged orientation of these residues is unknown, these residues must be excluded from the derivation of the anisotropic diffusion tensors.

### **5.3.2 Comparison of the Experimental Isotropic, Axially Symmetric, and Fully Anisotropic Models for Describing the Overall Rotational Diffusion of GB3**

The parameters that describe the overall rotational diffusion of GB3 determined from relaxation data ( $R_1$ ,  $R_2$ ,  $^{15}\text{N}\{^1\text{H}\}$  NOE) at 600 MHz assuming three different models of motion (isotropic overall rotational diffusion, rotational diffusion which has a single symmetry axis (i.e. is axially symmetric), and fully anisotropic rotational diffusion) are given in Table 5.3.1 along with statistics describing the fit of the model to the data. Of note here, the numbers in parentheses represent standard errors in the parameters arrived at by Monte Carlo simulation of synthetic parameter

sets about the  $\chi^2$  minimum using the method of  $\chi^2$  boundaries as described in <sup>117</sup>.

These standard errors represent 68.3% confidence intervals in the fit parameter, but should not be compared with the standard deviations in the distribution of this parameter for all residues in GB3. Three different theoretical predictions of this diffusion tensor, using hydrodynamic models are shown in the bottom three rows (the methods for these predictions and their agreement with the experimental values are discussed in section 5.4). The orientation of the axes of the derived diffusion tensors with respect to the structure of GB3 are shown in Figure 5.3.1.

**Table 5.3.1. Hydrodynamic characteristics of the GB3 domain derived from  $^{15}\text{N}$  relaxation data using various models of the overall tumbling and from hydrodynamics calculations**

| Model of overall motion             | $D_x^a$        | $D_y^a$        | $D_z^a$        | $\alpha^b$ | $\beta^b$  | $\gamma^b$  | $\tau_c^c$     | Anisotropy <sup>d</sup> | Rhombicity <sup>e</sup> | $\chi^2/\text{df}^f$ | P <sup>g</sup>     |
|-------------------------------------|----------------|----------------|----------------|------------|------------|-------------|----------------|-------------------------|-------------------------|----------------------|--------------------|
| Isotropic                           | 4.86<br>(0.04) | 4.86<br>(0.04) | 4.86<br>(0.04) | -          | -          | -           | 3.43<br>(0.03) | 1                       | 0                       | 102.5                | -                  |
| Axial symmetry                      | 4.45<br>(0.11) | 4.45<br>(0.11) | 6.07<br>(0.33) | 94<br>(7)  | 69<br>(12) | -           | 3.34<br>(0.11) | 1.37<br>(0.06)          | 0                       | 8.4                  | $4 \cdot 10^{-17}$ |
| Full anisotropy                     | 4.13<br>(0.24) | 4.60<br>(0.18) | 6.25<br>(0.34) | 85<br>(10) | 68<br>(7)  | 179<br>(14) | 3.34<br>(0.10) | 1.43<br>(0.09)          | 0.37<br>(0.24)          | 6.2                  | $5 \cdot 10^{-3}$  |
| Theoretical prediction <sup>h</sup> | 4.36           | 4.99           | 6.01           | 70         | 84         | 152         | 3.22           | 1.29                    | 0.71                    |                      |                    |
| Theoretical prediction <sup>i</sup> | 4.35           | 4.49           | 5.98           | 75         | 61         | 172         | 3.23           | 1.35                    | 0.13                    |                      |                    |
| Theoretical prediction <sup>j</sup> | 4.43           | 4.64           | 6.36           | 87         | 64         | 175         | 3.31           | 1.40                    | 0.17                    |                      |                    |

Numbers in the parentheses represent standard errors, arrived at by Monte Carlo simulation of synthetic parameter sets around the  $\chi^2$  minimum using the method of  $\chi^2$  boundaries as described in Press et al.

<sup>a</sup> Principal values (in  $10^7 \text{ s}^{-1}$ ) of the rotational diffusion tensor, ordered so that  $D_x \leq D_y \leq D_z$ .

<sup>b</sup> Euler angles  $\{\alpha, \beta, \gamma\}$  (in degrees) describe the orientation of the principal axes frame of the rotational diffusion tensor with respect to protein coordinate frame.

<sup>c</sup> Overall rotational correlation time (in ns) of the molecule,  $\tau_c = 1/[2 \text{Tr}(\underline{D})]$ .

<sup>d</sup> The degree of anisotropy of the diffusion tensor,  $2D_z/(D_x + D_y)$ .

<sup>e</sup> The rhombicity of the diffusion tensor,  $1.5(D_y - D_x)/[D_z - \frac{1}{2}(D_x + D_y)]$ .

<sup>f</sup> Residuals of the fit ( $\chi^2$ ) divided by the number of degrees of freedom.

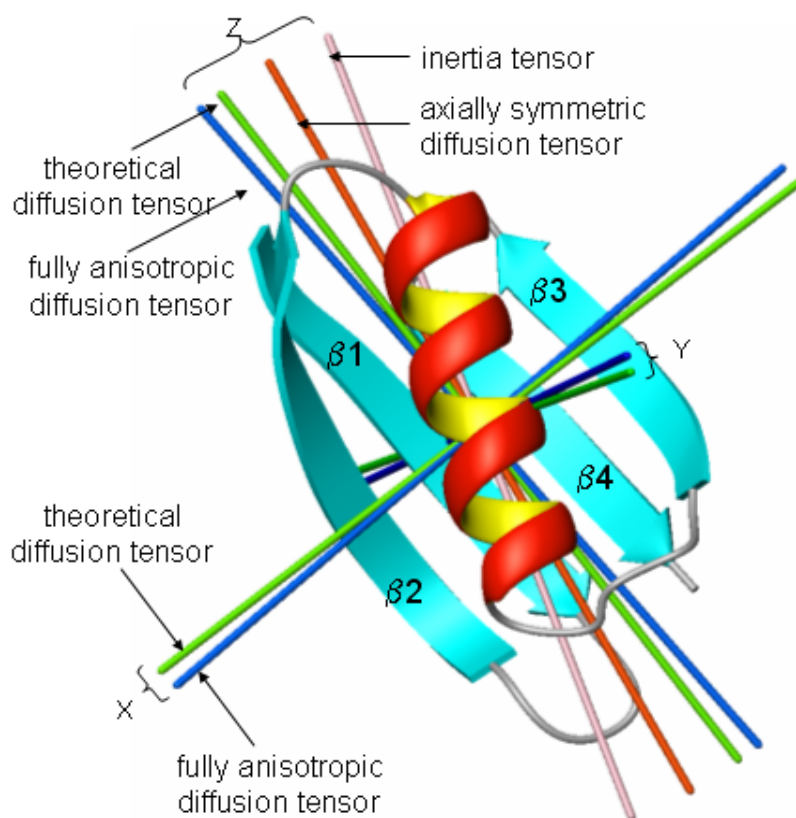
<sup>g</sup> Probability that the reduction in  $\chi^2$  (compared to the model in the row directly above it) could occur by chance. Both axially symmetric and fully anisotropic models are statistically a much better fit than the isotropic model.

<sup>h</sup> The results of hydrodynamic calculations using “dry” bead model, the bead radius was set to 1.4 Å.

<sup>i</sup> The results of hydrodynamic calculations using bead model and hydration shell, the bead radius was set to 1.0 Å with a hydration shell of width 1.3 Å.

<sup>j</sup> The results of hydrodynamic calculations using HYDRONMR program (García de la Torre et al., 2000); parameter  $a$  was set to 2.6 Å.

**Figure 5.3.2.** Ribbon representation of the tertiary structure of the GB3 domain, generated using MolMol (Koradi, J. Mol. Graph. 1996) with the orientation of the diffusion tensor axes (as indicated) obtained directly from  $^{15}\text{N}$  relaxation data for the axially symmetric (red) and fully anisotropic (blue) models and predicted theoretical (green) using HYDRONMR, along with the unique axis of the inertia tensor (pink). Atom coordinates are from the crystal structure (1IDG.pdb) (Derrick, J Mol. Biol. 1994). The orientations of all three diffusion tensors are similar within the experimental errors. The z-axis of the axially symmetric tensor makes an  $8^\circ$  angle with those for the fully anisotropic tensors, both measured and predicted using HYDRONMR. The difference in the orientation of the z-axes of the fully anisotropic and the theoretical tensor is  $3^\circ$ . All these z-axes are oriented approximately along the  $\alpha$ -helix axis: the tilt angle is  $23^\circ$ ,  $30^\circ$ , and  $28^\circ$ , for the axially symmetric, fully anisotropic, and the HYDRONMR-predicted tensors. Similar angles with respect to the unique axis of the inertia tensor are  $10^\circ$ ,  $18^\circ$ , and  $17^\circ$ ; this axis is tilted from the helix axis by  $15^\circ$ .



For the axially symmetric model, the rotational diffusion tensor of GB3 is characterized by  $D_{\parallel} / D_{\perp} = 1.37 \pm 0.06$  and  $\tau_c = 3.34 \pm 0.11$  ns. The orientation of the unique principal axis of this tensor with respect to the crystal structure (Fig.5.3.2) is characterized by Euler angles,  $\Phi = 94^\circ \pm 7^\circ$  and  $\Theta = 69^\circ \pm 12^\circ$ . The smaller the angle

between the axis of the diffusion tensor and the  $\alpha$ -helix axis (Fig.5.3.1) the more likely the observed elevation of the  $R_2$  values in this part of the protein are explained by the diffusional anisotropy of the molecule. Here the angle is only  $23^\circ$  for the axially symmetric tensor and  $30^\circ$  for the fully anisotropic tensor. The agreement between the experimental ( $\rho^{\text{exp}}$ ) and fit ( $\rho^{\text{calc}}$ , Eq 5.2.4) values of  $\rho$  as a function of the polar angle  $\theta$  between the NH vector and the axis of the axially symmetric diffusion tensor is shown in Fig.5.3.2a. The vertical spread of the data around the fitting curve, most pronounced near the maximum (at  $\theta \sim 90^\circ$ ), indicates that the actual diffusion of the molecule is not perfectly cylindrically symmetric, but is slightly rhombic.

The characteristics of the fully anisotropic diffusion tensor are very similar to those for the axially symmetric tensor (Table 5.3.1, Fig.5.3.2). In the case of full anisotropy the description of  $\rho$  in terms of one angle between the NH vector and the axis of the diffusion tensor is not sufficient, as the orientation (azimuthal angle  $\phi$ ) of the NH vector with respect to the plane containing the other two axes of the diffusion tensor also has an effect on the relaxation parameters. A more detailed assessment of the quality of the fit can therefore be obtained from the three-dimensional surface shown in Fig 5.3.3c and representing the theoretical values of  $\rho$  as a function of  $\theta$  and  $\phi$ . For a prolate fully anisotropic diffusion tensor, the shape of this surface displays the symmetry of  $\rho$ , i.e.  $\rho(\theta, \phi) = \rho(180^\circ - \theta, \phi) = \rho(\theta, -\phi) = \rho(\theta, 180^\circ - \phi) = \rho(\theta, \phi - 180^\circ)$ . Analogous to the height of the curve in Fig 5.3.3a, the elevation of this surface depends on the principal values  $\{D_x, D_y, D_z\}$  of the diffusion tensor. The difference in height between the maxima and the saddle points is proportional to  $(D_y - D_x)$  and

vanishes for axially symmetric diffusion. For the axially symmetric model, this whole surface is projected onto an area on the  $\rho(\theta)$  plot (compare Figs 5.3.3a,b). The upper and lower boundaries of  $\rho(\theta)$  are given by  $\phi = 0$  and  $\phi = \pi/2$  and correspond to the cases where the NH vector lies in the  $D_x$ - $D_z$  or  $D_y$ - $D_z$  plane. The gap between the two boundaries varies with the angle  $\theta$ ; it is negligible for  $\theta$  close to 0, it increases with the deviation of the NH vector from the z-axis and reaches maximum when  $\theta = \pi/2$ . The top points of these curves correspond to the two limiting cases of the NH vector parallel to the  $D_x$  axis ( $\theta = \pi/2, \phi=0$ , upper boundary) or along the  $D_y$  axis ( $\theta=\pi/2, \phi = \pi/2$ , lower boundary). The data points located in the “hills and valleys” on Fig. 5.3.3c are projected onto the space between the red and green lines in Fig. 5.3.3b when the surface is projected onto the  $\theta - \rho$  plane in the case of the axially symmetric model; this then explains the vertical spread in the data points around the fitting curve in Fig.5.3.3a. For an isotropic tensor the surface in Figure 5.3.3c is flat.

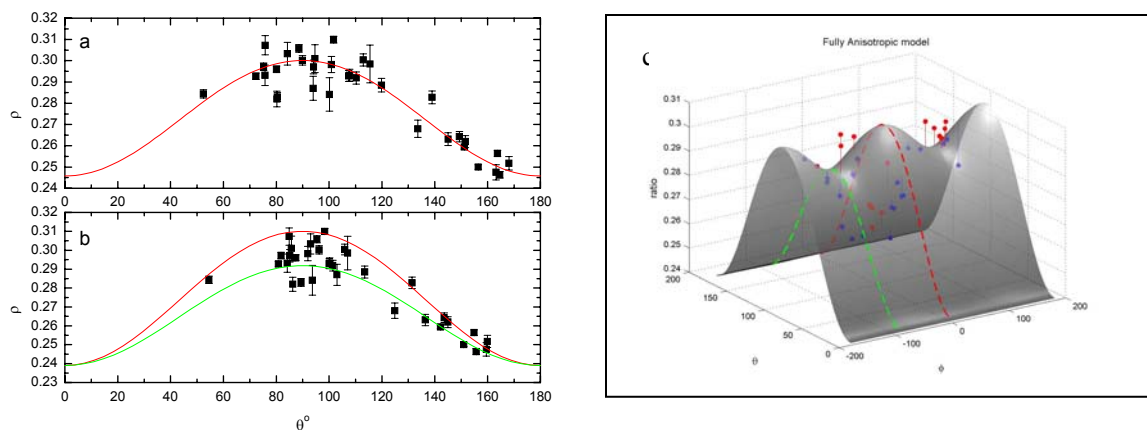
From the chi-square per degree of freedom ( $\chi^2/\text{df}$ ) of the fit (see Table 5.3.1), it can be seen that the axially symmetric and fully anisotropic tensors are a much better fit to the experimental data than the isotropic tensor ( $\chi^2/\text{df}_{\text{iso}}=102.5$ ,  $\chi^2/\text{df}_{\text{ax}}=8.4$ ,  $\chi^2/\text{df}_{\text{ani}}=6.2$ ), while the fully anisotropic tensor was only a slight improvement over the axially symmetric tensor. As discussed above, the actual diffusion of GB3 is slightly rhombic, but whether or not the additional two parameters necessary in describing fully anisotropic diffusion (compared to axially symmetric diffusion) improve the modeling of the data sufficiently to statistically justify their introduction remains in question. A statistical F-test was performed to evaluate whether this slight improvement in  $\chi^2/\text{df}$  is significant. The probability that



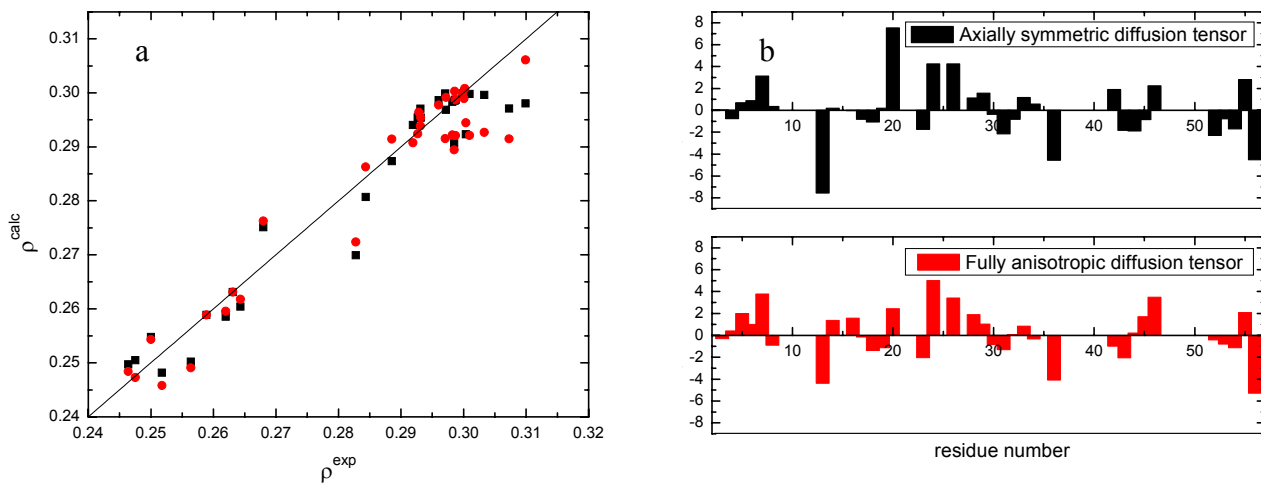
the fully anisotropic tensor is a better fit purely by chance is  $1/200$ . For F test statistics to be significant at the  $(1-\alpha)\%$  confidence level, this probability should be less than  $\alpha$ . Therefore, the fully anisotropic model is the best model to a 99.5% confidence level. However, the rhombicity of the fully anisotropic diffusion tensor is very small, and close to the minimum identifiable rhombicity (0.3 for relaxation rates with 2% uncertainty<sup>120</sup>). Therefore we concluded that the axially symmetric tensor and fully anisotropic tensor both provide approximately equal fits, so that to a good approximation, the diffusion of GB3 can be modeled by an axially symmetric diffusion tensor with the parameters in the second row of Table 5.3.1. This issue will be explored further in Chapter 7.



**Figure 5.3.3. Orientation dependence of  $\rho^{\text{exp}}$ .** In (a) and (b)  $\rho^{\text{exp}}$  (black squares) are compared to  $\rho^{\text{calc}}$  (smooth curves) using an axially symmetric (a) and fully anisotropic diffusion tensor that minimizes the corresponding  $\chi^2/\text{df}$  function (Eq. 4.2.1). (b). In (a) the red curve corresponds to Eq. 4.2.4 using the parameters in the second row of Table 4.3.1 In (b) are the curves representing the projection of the dependence on  $\phi$  and  $\theta$  in (c) onto the  $\theta - \rho$  plane. The red and green lines indicate the positions of the  $\phi=0^\circ$  (NH vector along  $D_y$ ) and  $\phi=90^\circ$  (NH vector along  $D_x$ ) as a function of  $\theta$ . These lines are close together for GB3, reflecting its small rhombicity. (c) shows the complete surface (in grey) that represents  $\rho^{\text{calc}}(\phi, \theta)$  calculated with the parameters in the third row of Table 4.3.1 for the fully anisotropic model. The red and blue points are  $\rho^{\text{exp}}$  in GB3 colored according to whether they lie above (red) or below (blue) the  $\rho^{\text{calc}}$  surface.



**Figure 5.3.4. (a) Correlation of  $\rho^{\text{exp}}$  and  $\rho^{\text{calc}}$  for the axially symmetric (black points) and fully anisotropic diffusion tensors (red points). (b) Difference between  $\rho^{\text{exp}}$  and  $\rho^{\text{calc}}$  divided by the error in  $\rho^{\text{exp}}$  ( $\sigma$ ) for each residue in GB3. The sum over all residues of the square of this difference is the function minimized in the program ROTDIF, Eq. 5.2.1.**



#### **5.4 Comparison with Predictions from Theoretical Hydrodynamic Models**

Having experimental data for the rotational diffusion tensor of the protein, we can now test if theoretical models are capable of reproducing these results. Theoretical prediction of the rotational properties of proteins in solution is complex, largely because it has to account for the unknown size and shape of the hydration shell formed by nearby water molecules moving together with the tumbling protein molecule. A detailed theoretical analysis should consider specific interactions between water molecules and protein atoms and the friction effects due to the roughness of the protein surface<sup>122</sup>. In addition, large-amplitude dynamics (e.g. of the loops and/or termini) can alter the shape of the molecule in time, in which case a rigid-body approximation is invalid. Here we considered several ways of theoretically predicting the overall rotational diffusion tensor of GB3, based on several representations of the protein's structure, with increasing levels of modeling sophistication.

##### **5.4.1 Simple Predictions based on the Stokes-Einstein-Debye Equation of Rotational Diffusion and Empirical Relations for Proteins.**

The rotational properties of a rigid rotor in a frictionless medium are characterized by its inertia tensor. The inertia tensor of GB3 was calculated using the coordinates of the heavy atoms from the crystal structure (1IDG.pdb). The inertia tensor of GB3 has normalized principal values of 1.80: 1.79: 1.00, indicating that the protein can be modeled as an axially symmetric rotor. The unique axis of the tensor is approximately parallel to the  $\alpha$ -helix axis as shown in Fig. 5.3.2.

The relationship between the inertia tensor of an object and its rotational diffusion tensor, relatively straightforward for rigid objects of simple shape, becomes more complex for realistic representations of a protein. A rough theoretical estimate of the principal components,  $D_i$  ( $i = x, y, \text{ or } z$ ), of the diffusion tensor and of the overall correlation time,  $\tau_c$ , of GB3 can be made assuming a Stokes-Einstein-Debye hydrodynamics model in which the protein is approximated as a rigid rotor in the shape of a sphere, cylinder, or prolate ellipsoid of revolution.

The simplest model is to represent a protein by a sphere. Using the Stokes-Einstein-Debye equation:  $\tau_c = \eta V / k_b T$ , we obtain  $\tau_c = 1.64$  ns. Here  $\eta$  is the solvent viscosity,  $T$  is temperature,  $k_b$  is the Boltzmann constant, and  $V$  is the volume of the molecule which we estimated from the molecular weight of the protein assuming that the specific volume is uniformly  $0.73 \text{ cm}^3/\text{g}$ .

The cylinder approximation using empirical relationships from the literature<sup>123</sup> resulted in  $D_{\parallel} / D_{\perp} = 1.45$  and  $\tau_c = 2.35$  ns, assuming solvent viscosity of 0.91 cpoise at 24°C. The sizes of the molecule in the relevant dimensions ( $27\text{\AA}$  in the  $z$  and  $16\text{\AA}$  in both  $x$  and  $y$ , the axial ratio 1.69) were obtained from the crystal structure.

For a prolate ellipsoid model, the ratio of the principal values of the diffusion tensor is approximately given by the empirical relationship  $D_{\parallel} / D_{\perp} = (I_{\perp} / I_{\parallel})^{1/\sqrt{2}}$ <sup>124</sup>, where  $I_{\parallel}$  and  $I_{\perp}$  are the principal components of the inertia tensor of the protein.

Using this model for GB3, we obtained  $D_{\parallel} / D_{\perp} = 1.51$  and  $\tau_c = 2.64$  ns under identical solvent conditions.

### 5.4.2 Modern Hydrodynamic Models

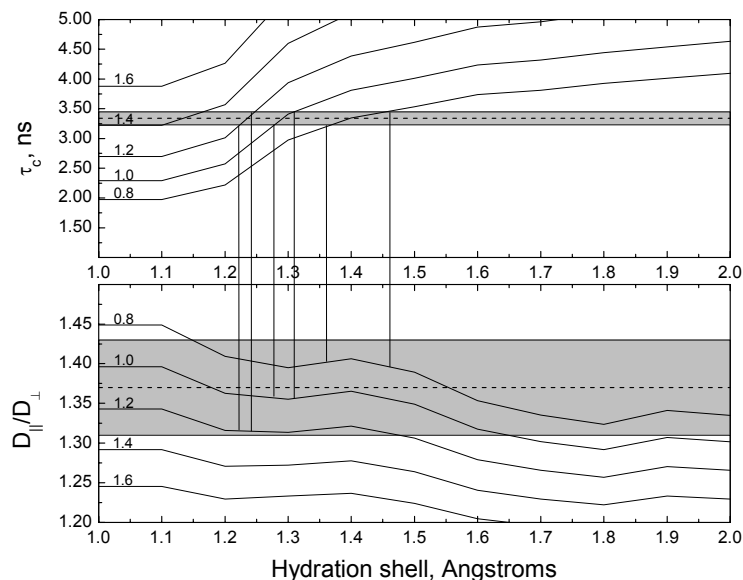
We then considered more sophisticated models that take into account atomic-resolution details of the shape of the protein and attempt to include the effect of the solvent. Since a detailed picture of the protein's interactions with the surrounding water molecules is not available, these interactions are modeled by including a hydration layer of uniform width that tumbles together with the protein. A more detailed theoretical analysis would consider specific interactions between water molecules and protein atoms<sup>125</sup> and the friction effects due to the fractal nature of the protein surface<sup>122</sup>.

Here we focused on two characteristics of the diffusion tensor: its anisotropy and the overall correlation time. We selected these parameters because of the opposite character of their dependence on the size of the hydration shell: adding a layer of water molecules will increase  $\tau_c$  (as the rotating body is now larger) and decrease the  $D_{\parallel}/D_{\perp}$  (as the hydration shell enclosed protein is more rounded than the protein alone). Therefore a simultaneous comparison of the predictions for both characteristics of the tensor could provide insights into the optimal settings for theoretical hydrodynamic models.

The so-called 'bead model' for prediction of the hydrodynamic properties of molecules approximates protein by a series of beads<sup>126-128</sup> placed at the coordinates of heavy atoms and with the bead size representing the average atomic radius. First we considered a "dry protein". For the B3 domain of protein G we could reproduce the experimentally obtained value of  $\tau_c$  for an atomic radius of 1.45 Å, and the experimentally obtained value of  $D_{\parallel}/D_{\perp}$  for a radius of 1.1 Å.

Theoretically, one average atomic radius should reproduce both experimental parameters of overall rotational diffusion. Therefore we conclude that the “dry protein” model is not adequate. We then included hydration shells of increasing thickness ( $0\text{ \AA}$  -  $5\text{ \AA}$ ) to the protein bead model to test if this could reproduce the values of both experimental parameters ( $D_{\parallel}/D_{\perp}$  and  $\tau_c$ ) for one bead size and one shell thickness. It turns out that several combinations of bead size and hydration shell thickness are consistent with the experimental values (see Fig. 5.4.1), given the experimental uncertainties. The optimal bead sizes ranged from  $0.8$  to  $1.2\text{ \AA}$  and the corresponding values of the shell thickness from  $1.5$  to  $1.2\text{ \AA}$ . These values are somewhat smaller than those typically assumed in hydrodynamic calculations—probably reflecting the incomplete modeling of the solvation of the protein by a hydration layer of uniform thickness. For example, it has been shown that the hydration layer surrounding a protein consists of two different “types” of water<sup>129</sup> — a few water molecules that remain tightly bound to the protein for long times (greater than the rotational correlation time of the molecule) and other water molecules still near the protein surface that experience much faster rotational and translational diffusion rates. The results of the “dry” bead model with the bead radius set to  $1.4\text{ \AA}$  and the results of the bead model with hydration shell with bead radius of  $1.0\text{ \AA}$  and hydration shell width of  $1.3\text{ \AA}$  are shown in Table 5.3.1 for comparison with experimental data.

**Figure 5.4.1.** Comparison of the measured characteristics,  $\tau_c$  and  $D_{\parallel}/D_{\perp}$ , of the diffusion tensor with the results of a hydrodynamic bead model calculations. Shown is the dependence of  $\tau_c$  (top) and  $D_{\parallel}/D_{\perp}$  (bottom) on the hydration shell thickness for various atom “bead” sizes, indicated by the corresponding numbers for each line. The dashed lines represent the experimental values of the diffusion tensor characteristics derived for the axially symmetric model, while the shaded areas represent their 68.3%-confidence region. The vertical bars mark the regions which are inside the experimental errors for both measured parameters ( $D_{\parallel}/D_{\perp}$  and  $\tau_c$ ) for a particular bead size.



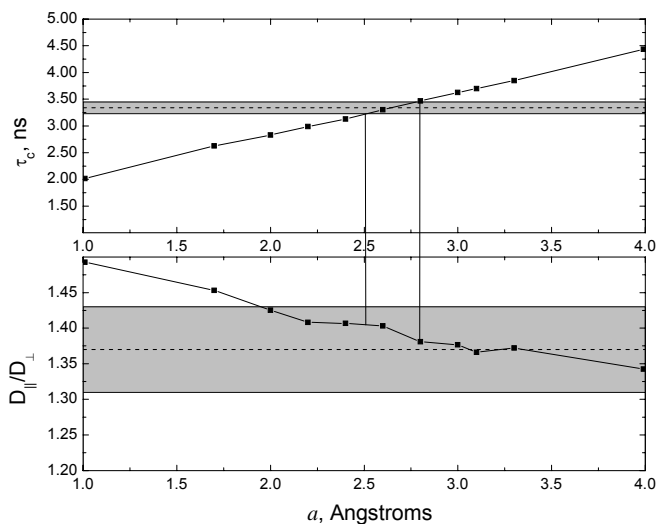
Another, more recent method for calculating surface effects of molecules in solution, uses a strategy known as shell modeling<sup>130,131</sup>, where the hydration effects are represented by a shell covering the surface of the protein. This model is characterized by a single parameter  $a$  that represents the sum of the thickness of the hydration shell and the average atomic van der Waals radius in the molecule. For the B3 domain of protein G we were able to reproduce the experimentally obtained values for both  $\tau_c$  and  $D_{\parallel}/D_{\perp}$  for an  $a$  between 2.5 Å and 2.8 Å (Fig. 5.4.2). As shown in Table 5.3.1 and Figure 5.3.1, both the principal values and the orientation of the calculated diffusion tensor are in remarkable agreement with the experimental data. Assuming an average atomic van der Waals radius of a heavy atom in the protein is about 1.5 Å, the hydration shell should have a thickness of between 1.0 Å and 1.3 Å.



This is generally consistent with the results of the bead model, where the hydration shell thickness was calculated to be between 1.2 Å and 1.5 Å.

Values of parameter  $a$  between 2.5 Å and 2.8 Å are consistent with the results obtained by de la Torre et al.<sup>130</sup> who calculated rotational diffusion tensors for a variety of  $a$  values for 15 proteins covering a range of molecular weights from 2.93 to 26.7 kDa. They found that in most cases experimental values of  $\tau_c$  were reproduced with values of  $a$  between 2 Å and 4 Å.

**Figure 5.4.2.** Comparison of the measured characteristics,  $\tau_c$  and  $D_{\parallel}/D_{\perp}$ , of the diffusion tensor with the results of the hydrodynamic shell model calculations. Shown is the dependence of  $\tau_c$  (top) and  $D_{\parallel}/D_{\perp}$  (bottom) on the parameter  $a$  (average van der Waals radius of the atoms in the molecule, plus the thickness of the hydration shell). The dashed lines, shaded regions, and vertical bars have the same meaning as in Figure 5.4.1.



## 5.5 Conclusions and Discussion

We determined experimentally the overall rotational diffusion tensor and correlation time of the GB3 protein from  $^{15}\text{N}$  relaxation rates ( $R_1$ ,  $R_2$ ,  $^{15}\text{N}\{^1\text{H}\}$  NOE) at 600 MHz. We found that to a good approximation, this protein can be modeled as a prolate axially symmetric (symmetric-top) rotor with the ratio of the rate of

reorientation about the fast axis to a perpendicular axis,  $D_{\parallel} / D_{\perp}$  of 1.37 and an overall correlation time,  $\tau_c$  of 3.34 ns. The rotational diffusion of the protein is only very slightly rhombic, with  $D_z / D_y = 1.36$  and  $D_y / D_x = 1.11$ , and the improvement in the fit of the relaxation data using this diffusion tensor was not found to have statistical significance compared to the fit using the axially symmetric diffusion tensor. We therefore conclude the rhombicity of overall rotational diffusion is negligible. This issue will be explored further in Chapter 7.

These experimental results were then used to evaluate the predictions of theoretical hydrodynamic models with varyingly sophisticated models for the shape/surface of the protein. The experimental values of  $D_{\parallel} / D_{\perp}$  and  $\tau_c$  were found to be in good agreement with a “wet” bead model for three combinations of bead size and hydration layer thickness, the most physically reasonable of which had a bead size of 1.0 Å and hydration shell thickness of 1.3 Å, however these values for these parameters are difficult to reconcile with expectations based on the van der Waals radius of Carbon (1.7 Å) and the diameter of a water molecule (mean van der Waals diameter from 2.82 Å to 3.2 Å). The experimental values of  $D_{\parallel} / D_{\perp}$  and  $\tau_c$  and the orientation of the axes of the diffusion tensor of GB3 were also found to be in good agreement with the predictions of the hydrodynamic shell model generated by the HYDRONMR program. We were able to reproduce the experimentally obtained values for both  $\tau_c$  and  $D_{\parallel}/D_{\perp}$  for an  $a$  between 2.5 Å and 2.8 Å, and the difference in the orientations of the z-axes of the experimental fully anisotropic tensor and the theoretical HYDRONMR tensor (with  $a = 2.8$  Å) is only 3°.



## Chapter 6: Local Motion in GB3.

### 6.1 Introduction

Many biological processes (e.g. enzyme catalysis<sup>132</sup>, allostery<sup>133</sup>, ligand recognition and binding<sup>16</sup>, and protein signaling<sup>134</sup>) involve motions on the  $\mu$ s-ms time scale. Motion between two stable conformational states that occurs on the  $\mu$ s-ms timescale in proteins (so-called conformational exchange motion) has been suggested to be important in catalysis (<sup>135</sup>,<sup>18,132</sup>) and may be rate-limiting in some examples of ligand binding (<sup>16</sup>). Much is known about such processes from kinetic studies of rates of conversion of substrates into products, and there is great interest in understanding how the conformational dynamics of the enzyme affects (or determines) these rates. Accurate identification of conformational exchange motions in proteins by NMR, a technique which has the unique ability to detect and characterize motions in proteins in solution, at a multitude of specific atomic sites, and over a range of timescales, is important for our understanding of the role these motions play in the biological function of proteins.

As described in Chapter 3, the  $^{15}\text{N}$  transverse relaxation rate ( $R_2$ ) is sensitive to exchange motion because the exchange between the two different magnetic environments of the conformational states contributes to dephasing of transverse coherence. Such motions therefore, lead to an increase in  $R_2$ , so that  $R_2$  is frequently written as the sum of  $R_{2\text{free}}$  and  $R_{\text{ex}}$ , where  $R_{2\text{free}}$  is the relaxation rate constant in the absence of exchange, and when the exchange is fast on the NMR timescale  $R_{\text{ex}}$  is related to the populations of the exchanging states ( $p_A$ ,  $p_B$ ) by:

$$R_{ex} \sim \frac{p_A p_B \Delta\omega^2}{k_{ex}}, \quad (6.1.1)$$

<sup>17</sup> where  $\Delta\omega$  is the difference in frequency between the states A and B and  $k_{ex}$  is the sum of the forward ( $k_{AB}$ ) and reverse ( $k_{BA}$ ) rate constants for the exchange.

Because of their potentially high content of biological information,  $R_{ex}$  contributions to  $R_2$  have been the focus of many novel NMR measurement experiments <sup>136-138</sup> and several relaxation studies <sup>139-142</sup>. In the previous chapter we described how rotational diffusion anisotropy led to elevation of the  $R_2$  values in the helix of the GB3 domain with respect to the  $R_2$  values in other parts of the protein. Such an elevation could be incorrectly identified as  $R_{ex}$  motion, and indeed this elevation of  $R_2$  rates in the helix of the GB1 domain was mistaken for conformational exchange <sup>118</sup>. Here we describe several ways of unambiguously identifying (or excluding)  $R_{ex}$  contributions to  $R_2$ . We applied these methods to the GB3 domain and determined that the only possible significant exchange contribution was in the residue Val39 located in the loop between the  $\alpha$ -helix and  $\beta$ -strand  $\beta_3$ . We also show that for GB3,  $R_{ex}$  values derived from a Lipari-Szabo “model-free” analysis of the local dynamics of the NH bonds depend dramatically on the model of overall rotational diffusion used in the analysis. Only by using the anisotropic tensor derived in Chapter 4 can the correct picture (significant  $R_{ex}$  for Val39 only) of the motions of the individual NH bonds be determined. This illustrates the necessity of a correct treatment of overall rotational diffusion to the derivation of parameters describing localized motion in proteins from NMR relaxation data.

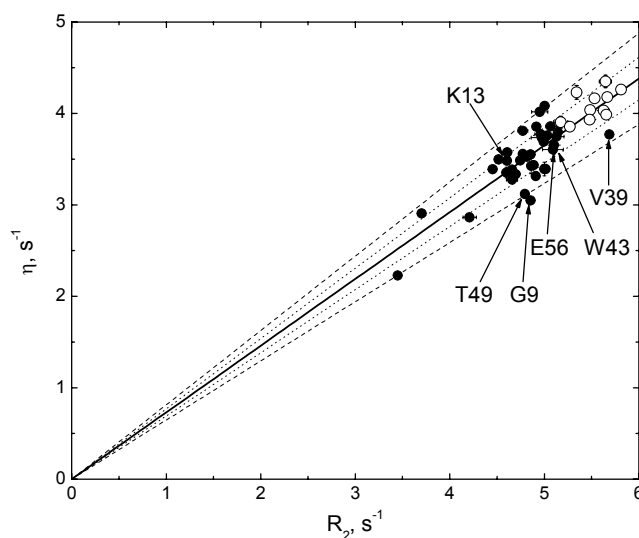
## **6.2 Combinations of Auto- and Cross-Correlation Relaxation Rates in GB3 Can Identify Conformational Exchange**

Methods for identifying conformational exchange motions in proteins can be broadly categorized as those that require knowledge of the protein structure and/or assumptions about the local or overall motions and those which do not. An example of the second type, a so-called “model-independent” approach is considered here, based on the direct comparison of  $R_2$  values with the transverse  $^{15}\text{N}$  CSA/dipolar cross-correlation rates,  $\eta_{xy}$ . Both  $\eta_{xy}$  and  $R_2'$  (see Eq. 2.5.20) depend on the same combination of spectral densities<sup>50,51</sup> (see Chapter 2 Eqs. 2.5.8 and 2.5.20), and thus sample motions on the same timescale, but unlike  $R_2'$ ,  $\eta_{xy}$  contains no contribution from conformational exchange. In the absence of conformational exchange,  $\eta_{xy}$  should scale linearly with  $R_2'$ , with the proportionality coefficient depending only on parameters reflecting local environment of the  $^{15}\text{N}$  nucleus: the NH bond length and the magnitude and orientation of the CSA tensor. Therefore, deviations from linearity of  $\eta_{xy}$  versus  $R_2'$  can be used to identify those sites involved in conformational exchange. As pointed out in<sup>50</sup>, this analysis does not require, hence is not biased by, any information on the protein structure, shape, tumbling rates, or preferred axes of rotation.

Figure 6.2.1 illustrates the linear relationship between  $\eta_{xy}$  and  $R_2'$  obtained by this comparison for GB3. Deviations of the data points from the average line may represent conformational exchange, and/or local variations in the  $^{15}\text{N}$  CSA and the angle  $\beta$  between the CSA and dipolar interaction<sup>50</sup>. While site-specific variations in the CSA and/or  $\beta$  will presumably cause the data to be distributed on both sides of the

average line (see Chapter 7 for a more complete discussion of the effect of site-specific CSA values), the  $R_{ex}$  contribution will increase  $R_2$  but not  $\eta_{xy}$  and therefore is expected to result in a horizontal shift of the data to the right.

**Figure 6.2.1. Linear agreement between experimental values of  $\eta_{xy}$  and  $R_2$ .** Residues in the  $\alpha$ -helix are shown as open circles, while the rest of the backbone amides are represented by solid circles. The fit line (solid) corresponds to a CSA of  $-160$  ppm assuming a  $\beta$  angle of  $20^\circ$  (or equivalently to a CSA of  $-175$  and  $22^\circ$ ), and  $r_{NH}=1.02\text{\AA}$ . The dashed lines represent the range of  $^{15}\text{N}$  CSA values (from  $-216$  ppm to  $-125$  ppm) observed in ubiquitin (here we assumed  $\beta=20^\circ$ ) while the dotted lines correspond to variations in  $\beta$  ( $20 \pm 5^\circ$ ) for CSA= $-160$  ppm. Note that all helix residues fall within region delimited by boundaries in variations in  $\beta$  and CSA and show no systematic shift to the right of the fit line, indicating that they are not involved in conformational exchange. The positions of Lys<sup>13</sup>, Val<sup>39</sup>, Trp<sup>43</sup> and Glu<sup>56</sup> are indicated. Of these residues, only Val<sup>39</sup> shows a significant shift to the right of the fit line, though it remains within the bounds of variations in CSAs measured in ubiquitin. Also indicated are positions of Gly<sup>9</sup> and Thr<sup>49</sup> that, together with Val<sup>39</sup>, are the most right-shifted residues.



The data points representing residues in the  $\alpha$ -helix all fall to the left of or within the error bars from the average line (Fig.6.2.1, open circles represent residues in the  $\alpha$ -helix). Since none of these residues is appreciably shifted to the right, we conclude that there is not conformational exchange in the helix. Residues that show a noticeable right shift, such as Gly<sup>9</sup>, Val<sup>39</sup>, and Thr<sup>49</sup>, are possible candidates for

conformational exchange, though the presence of a small shift alone is not a sufficient condition for determination of conformational exchange as it may reflect residue-specific variations in the  $^{15}\text{N}$  CSA or angle  $\beta$  (discussed in detail in Chapter 6) or some combination of these effects.

In order to remove the uncertainty associated with the site-specific variations in the magnitude and orientation of the  $^{15}\text{N}$  CSA tensor, we also compared the ratio,  $\eta_{xy}/\eta_z$ , of the transverse and longitudinal cross-correlation rates with that for the corresponding relaxation rates,  $R_2'/R_1'$ . As shown in <sup>101</sup>, in the absence of conformational exchange, the two ratios are equal within experimental errors. Therefore, such a comparison can be used as an indicator of the presence of conformational exchange which will increase  $R_2'/R_1'$  but should not affect  $\eta_{xy}/\eta_z$ . We therefore derived an “exchange-free” estimate of  $R_2'$ :

$$R_{2\text{free}}' = (\eta_{xy}/\eta_z) R_1'. \quad (6.2.1)$$

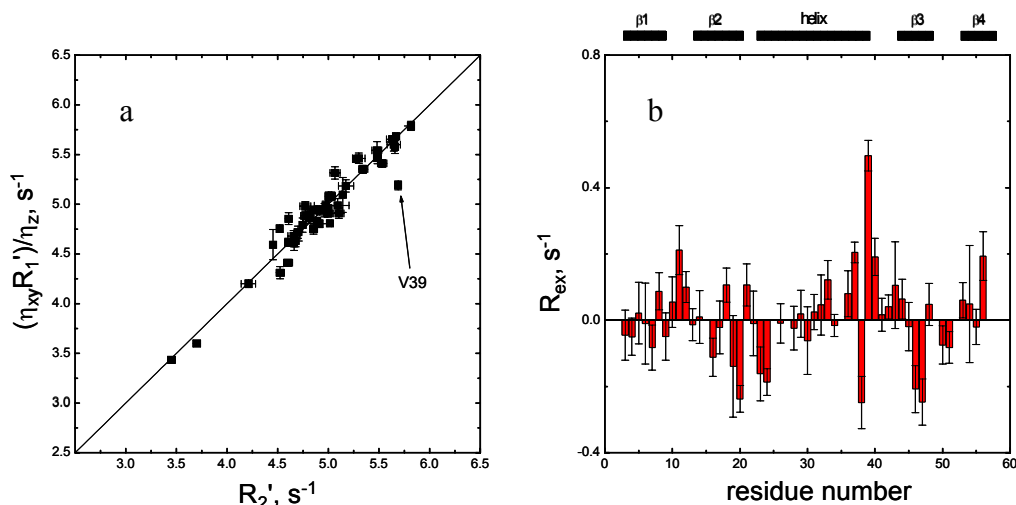
The remarkable agreement between the actual and exchange-free values of  $R_2'$  (Fig.6.2.2a) supports the conclusion that most of the backbone amides in GB3, except Val<sup>39</sup>, are not involved in any conformational exchange motions. The conformational exchange contribution can be estimated from these data as (see also <sup>101</sup>)

$$R_{\text{ex}} = R_2' - R_{2\text{free}}'. \quad (6.2.2)$$

This gives a  $R_{\text{ex}}$  estimate of  $0.50 \pm 0.05 \text{ s}^{-1}$  for Val<sup>39</sup> (Figure 6.2.2b).



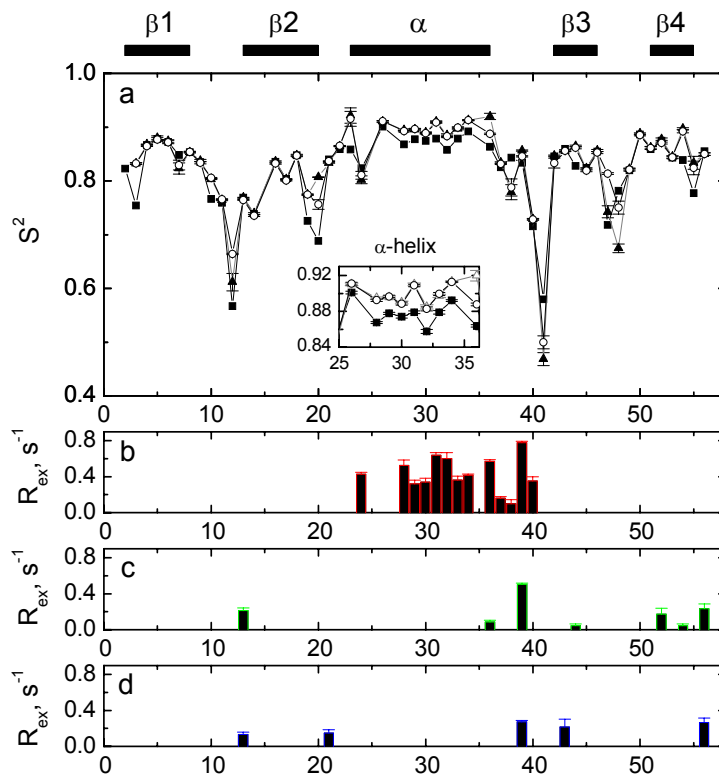
**Figure 6.2.2.** Comparison of the measured values of  $R_2'$  with their “exchange-free” estimates  $R_{2\text{free}}'$  (Eq. 6.2.1); the correlation coefficient between the two rates is  $r=0.96$  (0.97 if Val<sup>39</sup> is excluded). The shift of the data points to the right from the diagonal directly gives the  $R_{\text{ex}}$  values (shown in panel b). The spread of the data points around the diagonal and the small difference between the two ratios could be due to several factors, including measurement errors, the noncollinearity of the  $^{15}\text{N}$  CSA and  $^1\text{H}$ - $^{15}\text{N}$  dipolar interactions, and deviations of the  $^{15}\text{N}$  CSA tensor from axial symmetry.



### **6.3 Picture of Local Motion in GB3 is Markedly Dependent on the Model of Overall Rotational Diffusion**

The backbone microdynamic parameters ( $S^2$ ,  $\tau_{\text{loc}}$ ) and  $R_{\text{ex}}$  contributions (if any) were determined for GB3 using the following three models of the overall rotational diffusion tensor: isotropic, axially symmetric, and fully anisotropic. The overall correlation time for the isotropic model was optimized simultaneously with the model-free analysis (see <sup>143</sup>). The characteristics of the rotational diffusion tensor for anisotropic models were derived as described in Chapter 4 and then used to obtain residue-specific values of the microdynamic parameters. Values for  $S^2$  and  $R_{\text{ex}}$  for each residue are shown in Fig. 6.3.1

**Figure 6.3.1.** Comparison of the model-free parameters for backbone dynamics in GB3, determined using the three models of the overall tumbling. (a) Squared order parameters for isotropic (solid squares), axially symmetric (open circles), and anisotropic (solid triangles) models. Panels **b-d** depict the  $R_{ex}$  contributions to  $R_2$  obtained assuming (b) isotropic, (c) axially symmetric, and (d) anisotropic models of overall rotational diffusion. Insert in panel **a** is a blowup of the order parameters in the region of the  $\alpha$ -helix, to demonstrate the similarities and the differences in order parameters for the three models. The horizontal bars on the top of panel **a** indicate the elements of secondary structure.



All three models show a decrease in order parameters in the  $\beta 1/\beta 2$  loop and in the loop between the  $\alpha$ -helix and  $\beta 3$  indicating these are flexible regions. In the region between  $\beta 2$  and the  $\alpha$ -helix, and in the  $\beta 3/\beta 4$  loop there is a small decrease in order parameter, indicating that these regions are more flexible than the elements of secondary structure but less flexible than the other, more extended loops. All three models show elevated order parameters in the region of the  $\alpha$ -helix. The fully anisotropic and axially symmetric models predict slightly higher values for the order

parameters in this region than does the isotropic model. For most of the NH groups, the order parameters derived using the two anisotropic models are practically indistinguishable from each other; the exceptions are Leu<sup>12</sup>, Gly<sup>41</sup>, Asp<sup>47</sup>, and Ala<sup>48</sup> located in the flexible loops as well as Ala<sup>20</sup> and Asp<sup>36</sup> at the very edge of the  $\beta$ 2 strand and the  $\alpha$ -helix, respectively. Excluding Ala<sup>20</sup> and Asp<sup>36</sup> from the list of protein core residues for the fully anisotropic diffusion tensor analysis results in a 24% reduction in the  $\chi^2$  of the fit while the values of the derived parameters (e.g.,  $D_{\parallel}/D_{\perp}$ ,  $\tau_c$ , etc.) stay within their respective confidence limits. If these values are then used to calculate order parameters, there is no perceptible change except for Ala<sup>20</sup> and Asp<sup>36</sup> where the agreement between the axially symmetric and fully anisotropic models is significantly improved.

The values of local correlation time derived from these analyses varied from 0 to 56 ps for the majority of backbone amides, except for those (16 residues for the isotropic and 8 and 9 for axially and fully anisotropic diffusion tensors) where the extended model-free approach<sup>29</sup> was required. In the latter case, the correlation time for slow motions varied from 0.80 to 3.56 ns.

The most striking difference between the isotropic and anisotropic models is in the conformational exchange motions (Fig.6.3.1b-d). The isotropic model predicts conformational exchange in a stretch of 13 residues, Glu<sup>24</sup>, Lys<sup>28</sup>-Ala<sup>34</sup>, and Asp<sup>36</sup>-Asp<sup>40</sup>, covering the entire  $\alpha$ -helix (note that Thr<sup>25</sup>, Glu<sup>27</sup>, and Asn<sup>35</sup> were excluded due to signal overlap), while significantly fewer sites show this type of motion when rotational anisotropy is taken into account. Only for Val<sup>39</sup> do all three models agree, making it likely to exhibit conformational exchange. However, the value of  $R_{ex}$  for

this residue ( $0.78\text{ s}^{-1}$ ,  $0.51\text{ s}^{-1}$ , and  $0.28\text{ s}^{-1}$  for the isotropic, axially symmetric, and fully anisotropic models, respectively) seems to depend heavily on the model of overall motion. The other  $R_{\text{ex}}$  predictions of the anisotropic models, for Asp<sup>36</sup>, Thr<sup>44</sup>, Phe<sup>52</sup>, and Val<sup>54</sup> (axially symmetric) and Val<sup>21</sup>, Trp<sup>43</sup> (fully anisotropic) are not consistent between these two models nor with the isotropic model. Most of these  $R_{\text{ex}}$  values are too small (e.g. less than  $0.1\text{ s}^{-1}$  for Asp<sup>36</sup>, Thr<sup>44</sup>, Val<sup>54</sup> and  $< 0.17\text{ s}^{-1}$  for Val<sup>21</sup>, Phe<sup>52</sup>) to indicate real conformational exchange motion and probably represent errors in model selection.

Unlike the isotropic model which is essentially structure-independent, these anisotropic analyses all rely on information about the protein structure (specifically the orientation of the NH vectors with respect to the diffusion tensor) and, therefore, could be biased, if our derivation of the diffusion tensor (see Chapter 4) is in error. This could be particularly important for flexible regions of the protein, where the crystal structure might provide a snapshot rather than a representative orientation in solution, while a limited number (typically  $\sim 20$ ) of structures in the NMR ensemble might not provide proper conformational sampling. In addition, the results of model-free analysis could be biased by the underlying assumptions about the spectral density function and/or by the model-selection procedure. Therefore independent validation of these predictions is required based on methods (see section 6.2 and below) that do not directly require the knowledge of protein structure or any assumptions about models of motion.

## **6.4 Identification of Conformational Exchange Motion through the Field**

### **Dependence of $^{15}\text{N}$ Relaxation Rates**

Another model-independent method for differentiating between conformational exchange and the effects of overall rotational diffusion relies on the field dependence of the  $R_{\text{ex}}$  terms ( $R_{\text{ex}} \propto B_0^2$  see Eq. 6.1.1) in  $R_2$ . It is convenient to combine the available  $^{15}\text{N}$  relaxation parameters in the following form<sup>52,144</sup>:

$$R_J = 2R_2' - R_1' = (d^2 + c^2)4J(0) = \left( d^2 + \frac{\omega_N^2 \text{CSA}^2}{9} \right) 4J(0) \quad (6.4.1)$$

The field-independent terms,  $4d^2J(0)$  and  $((R_{\text{ex}}/\omega_N^2) + 2(\text{CSA}/3)^2J(0))$ , can be determined as the offset and slope, respectively, in the  $\omega^2$ -dependence of  $R_J$ , using relaxation measurements at multiple fields. This equation and information derived from it is discussed further in Chapter 7. As pointed out in<sup>52</sup>, this multiple-field method alone does not allow separate determination of the  $R_{\text{ex}}$  and CSA terms. However, it provides a direct relationship between these quantities, independent of any assumption about the overall or local motion, and therefore allows validation of the predictions of the model-free analysis. For example, for the data measured at two fields, indicated below by the subscripts '1' and '2', a little algebra on Eq. 6.4.1 gives the exchange contribution at a particular  $^{15}\text{N}$  Larmor frequency,  $\omega_N$ :

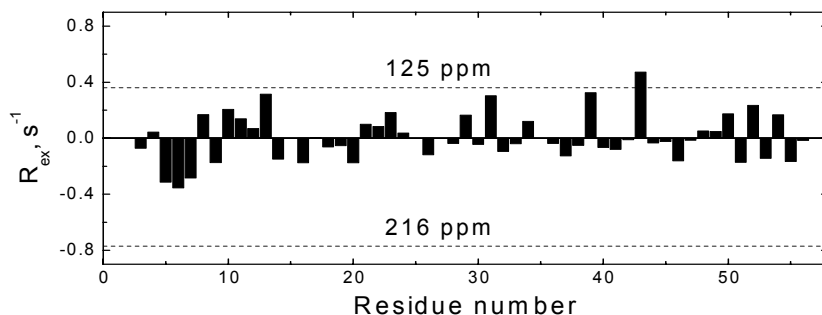
$$R_{\text{ex}} = \frac{\omega_N^2}{2} \left( \frac{R_{J,1} \left( 1 + \frac{\omega_{N,2}^2 \text{CSA}^2}{9d^2} \right) - R_{J,2} \left( 1 + \frac{\omega_{N,1}^2 \text{CSA}^2}{9d^2} \right)}{\omega_{N,1}^2 - \omega_{N,2}^2} \right) \quad (6.4.2)$$

This equation allows determination of the  $R_{\text{ex}}$  term as long as  $^{15}\text{N}$  CSA for a given group is known. In the absence of information about site-specific  $^{15}\text{N}$  CSA in GB3, it

is impossible to use this equation to accurately determine  $R_{ex}$  (see Chapter 7).

However, since a uniform  $^{15}\text{N}$  CSA value was assumed for the model-free analysis, Eq. 6.4.2 can be used to verify the predicted  $R_{ex}$  values. Figure 6.4.1a depicts the  $R_{ex}$  term derived from Eq.6.4.2 for an average  $^{15}\text{N}$  CSA of -160 ppm (as was assumed in the model free analysis). The figure also indicates the expected range of  $R_{ex}$  variations, assuming the range of  $^{15}\text{N}$  CSA values in GB3 is similar to that observed in ubiquitin<sup>51,52</sup>.

**Figure 6.4.1.**  $R_{ex}$  values calculated from the experimental data at 9.4 and 14.1 Tesla using Eq.4 and assuming -160 ppm as an average  $^{15}\text{N}$  CSA value. Dashed horizontal bars in **f** indicate the average levels of  $R_{ex}$  that one would obtain from Eq.4 for the indicated boundary CSA values.



Comparison of these  $R_{ex}$  values with the model-free results (Fig. 6.3.1b-d) for various models of overall diffusion described in the previous sections suggests that the isotropic model clearly gives false values of  $R_{ex}$  for all residues in the region of Ala<sup>23</sup>-Asp<sup>40</sup>, except Val<sup>39</sup>. As mentioned above, all overall models predicted conformational exchange contribution for Val<sup>39</sup>. A  $R_{ex}$  value slightly above the “noise” level, cf. Fig.6.3.1f, is obtained from Eq.4 for Lys<sup>13</sup>, consistent with the predictions from both anisotropic models, but not with the  $R_{ex}$  estimate from Eq.6.4.2. A  $R_{ex}$  term also derived from Eq.6.4.2 for Trp<sup>43</sup> seems to support the prediction from

the fully anisotropic model (Figs. 6.4.1 and 6.3.1d). This latter prediction, however, is probably not real, as it holds only for 3 out of the 24 NMR structures (see next section) and is not consistent with the  $R_{\text{ex}}$  values obtained from Eq.6.2.2 (Fig.6.2.2b). Lys<sup>31</sup> has a large predicted value of  $R_{\text{ex}}$  from Eq.6.4.2, and from all but one of the 24 NMR structures (Fig.6.5.1) but no  $R_{\text{ex}}$  is predicted for this residue when the orientation from the crystal structure is used and negligible  $R_{\text{ex}}$  is predicted for this residue using Eq.6.2.2 (Fig. 6.2.2).

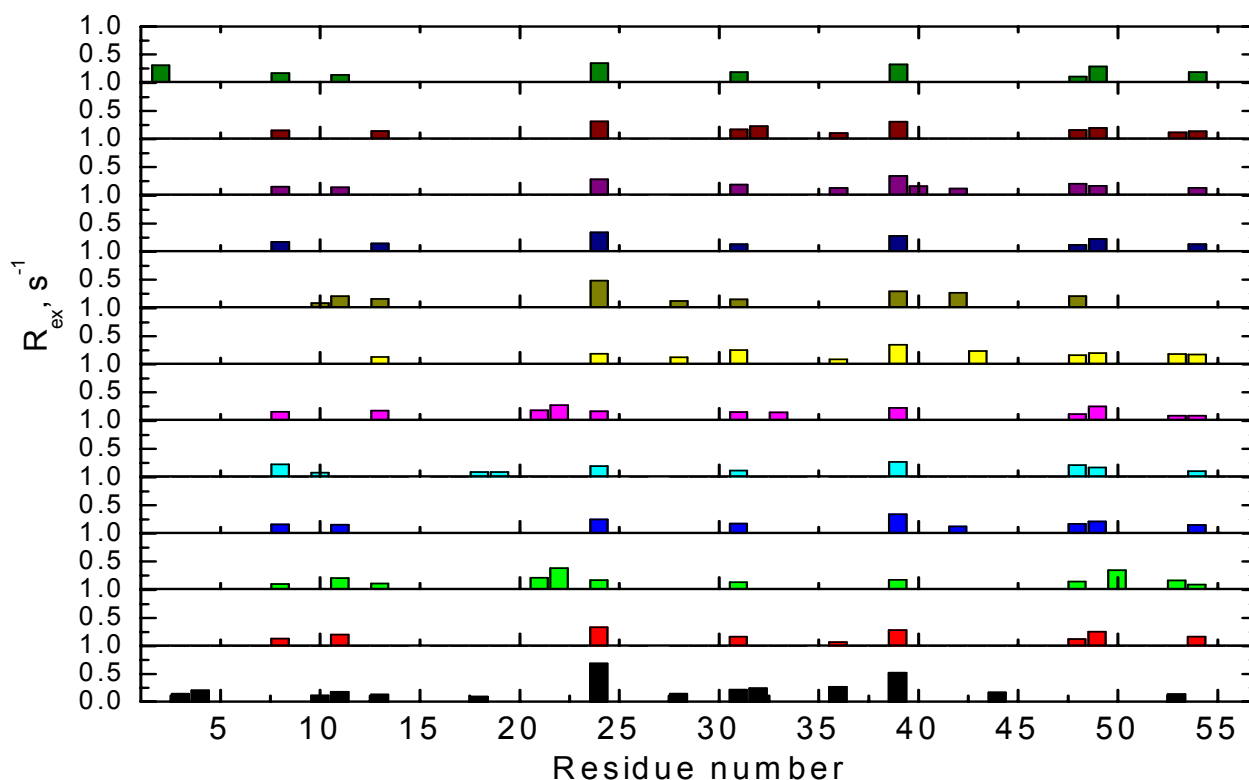
Note that, like the model-independent approaches presented in the previous section, the analysis based on Eq.6.4.2 does not require any information on protein structure or dynamics, and therefore is not biased by any assumption about the structure of protein molecule or its diffusion tensor.

## **6.5 Structural Dependence of Microdynamic Parameters and Conformational Exchange Motions**

Both the axial and fully anisotropic models (but not the isotropic model) suggest  $R_{\text{ex}}$  contributions for Lys<sup>13</sup>, and Glu<sup>56</sup>. Note that Val<sup>39</sup> is located in the middle of a flexible  $\alpha/\beta 3$  loop, Lys<sup>13</sup> is at the end of a flexible loop  $\beta 1/\beta 2$ , and Glu<sup>56</sup> is the C-terminal residue. The orientation of the NH vector for these residues might not be well defined, so the predicted  $R_{\text{ex}}$  values could be due to the orientation dependence of a particular “snapshot” of the NH bond orientation in the crystal structure rather than a real conformational exchange motion. To determine if such a bias due to orientation in the crystal structure exists, we performed similar analysis using a bundle of 24 NMR structures of GB3 (PDB file 2IGH)<sup>145</sup>. For all these structures, a

fully anisotropic model predicted conformational exchange for Val<sup>39</sup>, indicating that these motions are likely to be real (see also below). No  $R_{\text{ex}}$  contribution was found in Glu<sup>56</sup> for any of the 24 NMR structures, so the conformational exchange predicted for this residue was probably due to bias.  $R_{\text{ex}}$  terms (although small, less than 0.2 s<sup>-1</sup>) were obtained for Lys<sup>13</sup> in 16 out of the 24 NMR structures. This suggests that the predicted conformational exchange for Lys<sup>13</sup> is not likely to be due to bias caused by the use of the (single) crystal structure.

**Figure 6.5.1. Model-Free  $R_{\text{ex}}$  values based on the first 12 NMR structures from the 24 structure bundle (2IGH.pdb) assuming a fully anisotropic diffusion tensor. There is no predicted  $R_{\text{ex}}$  contribution for Glu56 for any of the structures. There are predictions of  $R_{\text{ex}}$  motion for all 12 structures for Val39, Glu24, and Lys31. However, if the NH orientations of Glu24 and Lys31 from the crystal structure (1IGD.pdb) are used, there is no such  $R_{\text{ex}}$  prediction. This indicates that the model-free  $R_{\text{ex}}$  predictions for these residues are not physical.**





## 6.6 Relaxation-Compensated CPMG Measurements Detect Conformational Exchange on Slower Timescales

Finally, we also conducted relaxation-compensated CPMG (Carr-Purcell-Meiboom-Gill) spin-echo measurements<sup>137,146</sup> using the pulse sequence of Loria and Palmer<sup>137</sup>, to investigate if there are any conformational exchange motions on a slower timescale, from 1 ms up to several ms, which the conventional  $R_2$  measurement would not sense (see Chapter 3). In these experiments, the adiabatic transverse relaxation rate for a particular frequency ( $\nu_{CPMG}$ ) of pulses in a spin-echo pulse train is given by:

$$R_{2(\nu_{CPMG})} = \varepsilon R_{IP} + (1 - \varepsilon) R_{AP} + R_{ex}, \quad (6.6.1)$$

where  $R_{IP}$  and  $R_{AP}$  are the transverse relaxation rate constants for in-phase and antiphase transverse coherences respectively (see Chapter 3), and  $\varepsilon$  is a constant (between 0 and 1) representing the population differences between in-phase and antiphase coherence as a function of the scalar coupling evolution of the Hamiltonian. For two exchanging sites, the rate constant for the exchange is given by:

$$R_{ex} = p_A p_B \frac{\Delta\omega^2}{k_{ex}} \left( 1 - \frac{4\nu_{CPMG}}{k_{ex}} \tanh\left(\frac{k_{ex}}{4\nu_{CPMG}}\right) \right), \quad (6.6.2)$$

in the fast exchange limit. Both  $\varepsilon$  and  $R_{ex}$  in equation 6.6.1 and 6.6.2 depend on

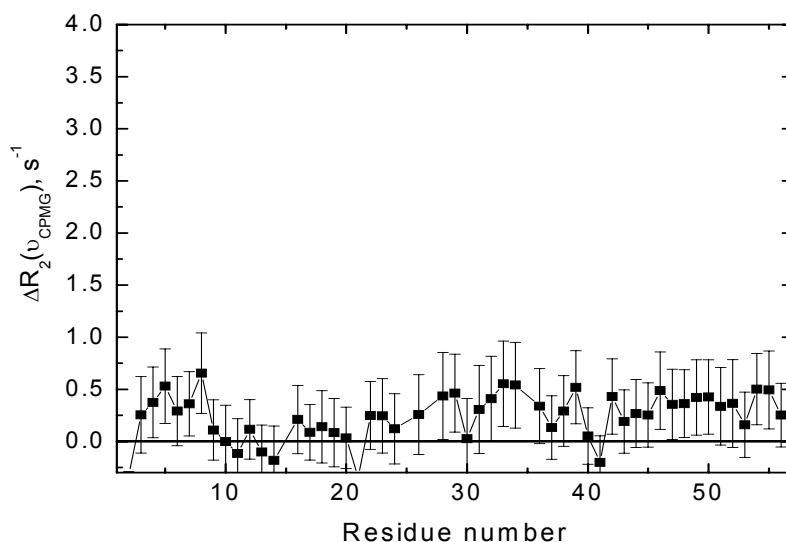
$\nu_{CPMG}$ , however, in the limit that  $\tau_{CPMG} = \frac{1}{\nu_{CPMG}} < \frac{1}{4J_{NH}}$ ,  $\varepsilon \approx 1$ , and if N and H are

decoupled throughout the experiment,  $\varepsilon = 1$ . In our experiment, the rate constants for in-phase and antiphase coherences are explicitly averaged (exactly as in the longitudinal cross-correlation rate measurement scheme presented in Chapter 4), so

that  $\varepsilon = 0.5$  for all  $\nu_{CPMG}$ . The difference in two values of  $R_2(\nu_{CPMG})$  is then indicative of an exchange contribution on a timescale between the two values of  $\tau_{CPMG}$ .

We measured  $R_2(\nu_{CPMG})$  with  $\tau_{CPMG}$  from 1ms up to 8ms in GB3. The difference in these rates  $\Delta R_2(\nu_{CPMG})$ , for all residues in GB3 are shown in Figure 6.6.1 and their values are close to 0 throughout the protein. This indicates no significant conformational exchange motion on the 1ms to 8ms timescale for those backbone amides for which 0 is within the experimental uncertainty of this measurement. Of note, there are five small stretches of residues in GB3 (4-5, 7-8, 33-34, 46-50, 54-55) and a few other residues (29, 39, 42) where  $\Delta R_2=0$  is not within the standard error (68% confidence interval) and for all of these residues the non-zero value of  $\Delta R_2$  is positive, and the distribution of  $\Delta R_2$  is not centered about zero. This could indicate small conformational exchange contributions on the ms timescale for these residues, or an underestimation of the experimental errors in the  $R_2$  rates.

**Figure 6.6.1. Conformational exchange in GB3, determined using the relaxation compensated CPMG sequence.**  $\Delta R_2(\nu_{CPMG})=R_2(\tau_{CPMG} = 8 \text{ ms})-R_2(\tau_{CPMG} = 1 \text{ ms})$  at 600 MHz is plotted versus residue number. There is little significant exchange in GB3 on this timescale (also, see text).



## **6.7 Conclusions**

Our analysis of model-free parameters describing the motion of NH bonds in the backbone of GB3 from  $^{15}\text{N}$  relaxation data shows that the isotropic and anisotropic models of the overall tumbling result in markedly different pictures of local motion; the main difference is in the interpretation of the elevated  $R_2$  values in the  $\alpha$ -helix: the isotropic model results in conformational exchange throughout the entire helix, whereas no exchange is predicted by anisotropic models that place the longitudinal axis of diffusion tensor almost parallel to the helix axis. Both axially symmetric and fully anisotropic models for the overall motion fit the experimental data significantly better than does the isotropic model.

Since the results of the anisotropic models could be biased by the available structural information, additional, model-independent methods for identification of exchange motions are required which do not rely on knowledge of protein structure or assumptions about its dynamics in order to distinguish the correct picture of motion. Three such methods are applied here to differentiate between the effects of conformational exchange and rotational anisotropy: a comparison of the CSA/dipolar cross-correlation rates ( $\eta_{xy}$ ,  $\eta_z$ ) with relaxation rates ( $R_2$ ,  $R_1$ ), the estimation of  $R_{\text{ex}}$  terms from  $^{15}\text{N}$  relaxation data at two fields, and relaxation-compensated CPMG measurement of exchange from experiments with different values of  $\nu_{\text{CPMG}}$  to identify exchange on longer timescales (1-8 ms). These methods are (1) sensitive to conformational exchange, and (2) do not require knowledge of protein structure or (3) any assumption about the spectral density function, and therefore can be used to

either identify potential sites for conformational exchange or verify the absence thereof. These analyses provide no indication of conformational exchange in the helix, consistent with the predictions of the model-free analysis using anisotropic models for the overall diffusion. This confirms our derived values of the parameters describing the overall diffusion tensor of GB3 (Chapter 4).

Furthermore, subsequent independent studies of GB3 have show that the order parameters arrived at here, using the axially symmetric diffusion tensor, are in excellent agreement with those derived from residual dipolar couplings measured in five alignment media<sup>147,14,148,149</sup>. This serves as an additional confirmation that there is no significant conformational exchange motion and further supports the axially symmetric model of the overall rotational diffusion of GB3 (Chapter 4).

## Chapter 7: $^{15}\text{N}$ Chemical Shielding Tensors in GB3

### 7.1 CSAs from NMR Relaxation Data; a Discussion of Current Techniques and Measurements

#### 7.1.1 Motivation

There are a variety of recent techniques for measurement of the individual chemical shift anisotropies of backbone  $^{15}\text{N}$  atoms in proteins in solution from auto and cross-correlation relaxation rates<sup>50-53 54-56</sup>. Thus far, residue specific solution measurements of the  $^{15}\text{N}$  CSA from relaxation measurements have been reported in three proteins: ubiquitin<sup>50-56</sup>, ribonuclease H, and a small alpha helical protein, C12A-p8<sup>MTCP 158-60</sup>.

Measurements in ubiquitin<sup>51,52</sup> revealed a range of site-specific backbone  $^{15}\text{N}$  CSA values, from approximately  $-120$  to  $-220$  ppm, with a mean of  $-157$  ppm and a standard deviation of 19 ppm. This range includes data for both conformationally well-defined amides and those located in the flexible regions. The angle between the unique axis of the  $^{15}\text{N}$  CST and the NH bond was found to vary from  $6^\circ$  to  $26^\circ$ , with the mean of  $15.7^\circ$  and standard deviation (std) of  $5^\circ$ <sup>51,52</sup>. These findings were confirmed by independent relaxation studies in ubiquitin<sup>57</sup> based on a model-free form of the spectral density function. A higher in absolute value average CSA of  $-173$  ppm (converted to an NH distance of  $1.02\text{\AA}$ ) with site-to-site variation (see section 7.4.5) of up to  $\pm 17$  ppm was derived from shieldings in peak positions in weakly aligned solutions of ubiquitin<sup>45</sup>, while recent MAS studies<sup>150</sup> of aligned ubiquitin in a similar medium yielded  $-162.0 \pm 4.3$  ppm for  $^{15}\text{N}$  CSA and  $18.6^\circ \pm 0.5^\circ$  for the angle, in agreement with those from previous  $^{15}\text{N}$  relaxation data<sup>51,52</sup>. A recent

study<sup>54</sup> combining new experimental measurements in ubiquitin with the literature data<sup>51,57</sup>, resulted in an even higher mean <sup>15</sup>N CSA of -179.6 ppm (converted to NH distance of 1.02 Å) and a very low CSA variability ( $\Lambda = \pm 5.3$  ppm, with upper limit  $\pm 9.4$  ppm at 95% confidence). However, the results of another recent study based on a combination of fourteen auto- and cross-correlation rates in ubiquitin<sup>56</sup> agree with the earlier data, and give average CSAs ranging from -146.4 to -164.0 ppm and the angles from 17.5-18.9°, depending on the choice of local motional model, with standard deviations from 10.1 to 13.7 ppm, and the site-to-site variability,  $\Lambda$ , ranging from 7.8 to 10.5 ppm, depending on the model of local motion.

A similar range of site-specific <sup>15</sup>N CSA values (-129 to -213 ppm) was reported for ribonuclease H<sup>53</sup>, although with a somewhat different mean (-172 ppm), of a selection of well-ordered amides. For this subset of residues, the site-to-site variability in CSA was estimated to be  $\pm 5.5$  ppm (upper limit  $\pm 9.6$  ppm at 95% confidence), assuming a Gaussian distribution for the <sup>15</sup>N CSA. This number is relatively small, given the ~30 ppm range of variation in the isotropic chemical shifts, and could be a result of the limited experimental precision in the CSA data, as the experimental uncertainties ( $\pm 13$  ppm) in the individual <sup>15</sup>N CSA values in that paper are noticeably bigger than the reported variability. Also, the analysis of such a restricted subset of amides (these selected “well ordered” amides account for only about 50% of the backbone amides in ribonuclease H) clearly precludes a definitive conclusion about the true range of variation in the CSA.

In the small alpha helical protein, C12A-p8<sup>MTCP1</sup><sup>58-60</sup>, an analysis of  $R_1$ ,  $R_2$ , and NOE rates at five field strengths found a mean CSA of -164 ppm with no

appreciable variability in the CSA above the uncertainty, though it should be noted that in this analysis all residue-specific field dependence was interpreted as due to conformational exchange motion (i.e. using Eq. 5.4.2, these contributions were subtracted out of  $R_2$  prior to the CSA analysis). Such interpretation erroneously attributes any actual variability in the CSA to site-specific variability in the exchange contribution to  $R_2$  and thus no conclusions can be drawn about the true range of variation in the  $^{15}\text{N}$  CSA from this analysis.

In all of these studies, the precision of the available relaxation data has been a limiting factor in determining if the distribution of  $^{15}\text{N}$  CSA's measured in solution for proteins agrees with that which has been reported from solid-state NMR studies on peptides. Specifically in question is whether or not there is significant variability in the  $^{15}\text{N}$  CSA from residue-to-residue (or site-to-site) within proteins and if this variability can be related to protein structure and/or chemistry. Additional studies of the distributions of  $^{15}\text{N}$  CSAs in other proteins are required to answer this question

### **7.1.2 Discussion of Multifield Analyses of $^{15}\text{N}$ Relaxation Rates**

When NMR relaxation measurements are made on multiple spectrometers, there is always the risk that small differences in sample conditions (e.g. in temperature or pH) or in experimental conditions (e.g. in pulse miscalibration or water suppression) or hardware will cause the measurements at multiple fields to be seemingly inconsistent with one another. Additionally, if the majority of residues are involved in conformational exchange motions,  $\tau_c$  estimated from  $\rho$  ratios (as described in Chapter 4) at different fields will seem to depend on the field strength. However, even for proteins where precautions were taken to keep sample,

experimental, and hardware conditions at multiple fields identical, and where conformational exchange was unlikely, inconsistencies in microdynamic parameters at multiple field strengths have been noted.

The problem of fitting relaxation data measured at several fields to the Lipari-Szabo form of the spectral density function was previously noted by Farrow et al.<sup>78</sup>, who observed that order parameters obtained from fitting relaxation data measured at several field strengths have low precision (though they should, in principle, be more accurate than order parameters obtained from data at one field strength) due to poor fits of multi-field data to a LS spectral density function. Other examples of discrepancies in the LS parameters derived from relaxation measurements at several fields can be found elsewhere<sup>34,44,151</sup> and have been attributed to various hypotheses such as time-diapason of intermolecular motions, additional ps-ns “floppiness” of proteins, and a failure of the LS spectral density function to accurately describe motions in proteins. Notably, Tjandra et al. have observed that using a value for the <sup>15</sup>N CSA of -170 ppm provided a better fit of LS spectral densities to relaxation rates in perdeuterated HIV protease measured at 360 and 600 MHz.

The chemical shielding anisotropy contribution to spin-relaxation is a field dependent effect. In conventional analyses of <sup>15</sup>N relaxation data this contribution is assumed to be uniform for all amides in the polypeptide backbone of the protein under study. It has previously been shown through computer simulation that significant variability in the <sup>15</sup>N CSA from residue to residue could lead to significant discrepancies in dynamic parameters at multiple fields<sup>100</sup>. However, an experimental verification of the variability in the <sup>15</sup>N CSA as the cause of discrepancy in dynamic



parameters estimated from  $^{15}\text{N}$  relaxation measurements at multiple fields has been missing.

### **7.1.3 Agreement of Diffusion Tensor of GB3 Derived from Auto and Cross-Correlation Relaxation Rates at Five Spectrometer Fields as Evidence that Relaxation Experiments Have Similar Conditions.**

For this analysis, the transverse ( $R_2$ ) and longitudinal ( $R_1$ )  $^{15}\text{N}$  relaxation rates and the steady-state  $^{15}\text{N}\{^1\text{H}\}$  NOEs in GB3 were measured at five magnetic fields, 9.4, 11.7, 14.1, 16.4 and 18.8 T of the same GB3 sample described in Chapter 3. The transverse ( $\eta_{xy}$ ) and longitudinal ( $\eta_z$ )  $^{15}\text{N}$  CSA/dipolar cross-correlation measurements were performed at four fields (9.4, 11.7, 14.1, and 18.8 T) for  $\eta_{xy}$  and at three fields (9.4, 11.7, and 14.1 T) for  $\eta_z$ . Fifty out of fifty-five amides were analyzed; residues Glu15, Thr25, Glu27, and Asn35 were excluded because of signal overlap and Val39 due to conformational exchange<sup>152</sup>. Gln2 was excluded from LS analyses since the atom coordinates for this residue (which is a mutation in our sample relative to the wild-type GB3 protein) were not available from the crystal structure.

There is a difference between these measurements and those analyzed in Chapters 4-6 in that the errors in the rates discussed previously were estimated by integrating regions of spectra containing no cross peaks, whereas here they were estimated from repeated (quadruplicate) measurements, using the method of<sup>97</sup>. In both cases the errors in the rates were estimated using a Monte Carlo simulation of 500 experimental data sets per residue and assuming a normal distribution of experimental errors in peak intensities. The experimental errors in relaxation rates using the repeated measurements were generally 2 to 10 times larger than those

estimated from integration of the spectral noise and were around 1% on average: 1.16%, 0.83%, 1.43%, 1.09%, and 1.37% for  $R_1$ ; 1.21%, 1.21%, 1.33%, 0.96%, and 1.30% for  $R_2$ , and 1.13%, 1.14%, 1.05%, 1.00%, and 1.06% for NOE values measured at 9.4, 11.7, 14.1, 16.4, and 18.8 Tesla, respectively. The average errors in  $\eta_{xy}$  were 1.37%, 1.50%, 1.67%, and 1.47% at 9.4, 11.7, 14.1 and 18.8 T, respectively; the errors in  $\eta_z$  were 1.27%, 1.16%, and 1.52% at 9.4, 11.7, and 14.1 T.

The axially symmetric diffusion tensors as derived from the relaxation measurements at each field individually are shown in Table 7.1.1 along with  $\chi^2/\text{df}$  describing the statistical goodness of fit. Isotropic and fully anisotropic diffusion tensors were also determined at each field strength, though in all cases the axially symmetric diffusion tensor was a sufficiently good fit to data as determined by F-test comparison with the isotropic and fully anisotropic models (see Chapter 4). The general agreement of the parameters describing the overall diffusion for all of the measurements indicates that there are no substantial experimental variations even though the measurements were made on different probes in different countries and separated in time by several months. Axially symmetric diffusion tensors derived from cross-correlation rate measurements ( $\eta_{xy}$ ,  $\eta_z$ ) as described in <sup>101</sup> are shown for comparison.

**Table 7.1.1 (a) Characteristics of the overall rotational diffusion tensor of GB3 derived from  $^{15}\text{N}$  relaxation data at five magnetic fields.** The NH vectors for this analysis were taken from the original crystal structure of GB3 (PDB entry 1IGD.pdb (Derrick and Wigley, 1994)). Also shown are the diffusion tensors derived from the CCR's  $\eta_{xy}$  and  $\eta_z$ .

| Magnetic field (Tesla)                            | $^1\text{H}$ resonance frequency (MHz) | $D_{\perp}^{\text{a}}$ ( $10^7 \text{ s}^{-1}$ ) | $D_{\parallel}^{\text{a}}$ ( $10^7 \text{ s}^{-1}$ ) | $\Phi^{\text{o b}}$ | $\Theta^{\text{o b}}$ | $\tau_{\text{c}}^{\text{c}}$ (ns) | Anisotropy <sup>d</sup> | $\chi^2/\text{df}^{\text{e}}$ | $P^{\text{f}}$     |
|---|--|--|--|---------------------|-----------------------|-----------------------------------|-------------------------|-------------------------------|--------------------|
| From auto- and cross-relaxation rate measurements |  |  |  |                     |                       |                                   |                         |                               |                    |
| 9.4   | 400                                    | 4.40(0.19)                                       | 6.13(0.62)   | 89(18)              | 66(23)                | 3.35(0.20)                        | 1.39(0.13)              | 0.64                          | $6 \cdot 10^{-11}$ |
| 11.7  | 500                                    | 4.45(0.31)                                       | 6.20(1.12)   | 95(15)              | 68(19)                | 3.31(0.32)                        | 1.39(0.24)              | 0.69                          | $4 \cdot 10^{-13}$ |
| 14.1  | 600                                    | 4.45(0.15)                                       | 6.05(0.44)   | 90(8)               | 70(10)                | 3.34(0.14)                        | 1.36(0.09)              | 0.72                          | $2 \cdot 10^{-13}$ |
| 16.4  | 700                                    | 4.44(0.14)                                       | 6.24(0.41)   | 99(7)               | 63(11)                | 3.31(0.13)                        | 1.41(0.08)              | 0.88                          | $6 \cdot 10^{-19}$ |
| 18.8  | 800                                    | 4.46(0.08)                                       | 6.15(0.27)   | 100(7)              | 67(10)                | 3.32(0.08)                        | 1.38(0.06)              | 0.74                          | $3 \cdot 10^{-14}$ |
| Averaged tensor                                   |  | 4.44   | 6.14   | 99                  | 66                    | 3.33                              | 1.38                    |                               |                    |
| Global-fit tensor                                 |  | 4.44   | 6.14   | 95                  | 66                    | 3.33                              | 1.38                    | 0.72                          | $6 \cdot 10^{-15}$ |
| From cross-correlation rate measurements          |  |  |  |                     |                       |                                   |                         |                               |                    |
| 9.4   | 400                                    | 4.50(0.16)                                       | 6.00(0.52)   | 101(9)              | 77(13)                | 3.33(0.16)                        | 1.33(0.11)              | 0.66                          | $9 \cdot 10^{-11}$ |
| 11.7  | 500                                    | 4.38(0.12)                                       | 6.14(0.40)   | 90(6)               | 59(9)                 | 3.36(0.12)                        | 1.40(0.08)              | 0.96                          | $1 \cdot 10^{-12}$ |
| 14.1  | 600                                    | 4.40(0.06)                                       | 6.20(0.19)   | 93(4)               | 65(6)                 | 3.33(0.06)                        | 1.41(0.04)              | 0.51                          | $3 \cdot 10^{-17}$ |

Numbers in the parentheses represent standard errors.

<sup>a</sup> Principal values of the rotational diffusion tensor.

<sup>b</sup> Polar and azimuthal angles  $\{\Theta, \Phi\}$  (in degrees) describe the orientation of the diffusion tensor axis with respect to protein coordinate frame.

<sup>c</sup> Overall rotational correlation time of the molecule,  $\tau_{\text{c}} = 1/[2 \text{tr}(\underline{D})]$ .

<sup>d</sup> The degree of anisotropy of the diffusion tensor,  $D_{\parallel}/D_{\perp}$ .

<sup>e</sup> Residuals of the fit divided by the number of degrees of freedom

<sup>f</sup> The probability that the reduction in the  $\chi^2$  compared to the isotropic diffusion model occurred by chance.

**Table 7.1.2. (b) Characteristics of the overall rotational diffusion tensor of GB3 derived from  $^{15}\text{N}$  relaxation data at different magnetic fields using two different residual dipolar coupling refined structures of GB3 (PDB entry 1P7E.pdb and 1P7F.pdb (Ulmer and Bax, 2003)).**

| Resonance frequency | $D_{\perp}^a$ | $D_{\parallel}^a$ | $\Phi^b$ | $\Theta^b$ | $\tau_c^c$ | Anisotropy <sup>d</sup> | $\chi^2/\text{df}^e$ |
|---------------------|---------------|-------------------|----------|------------|------------|-------------------------|----------------------|
| <u>1P7E.pdb:</u>    |               |                   |          |            |            |                         |                      |
| 400 MHz             | 4.49(0.38)    | 6.15(1.42)        | 106(21)  | 63(22)     | 3.31(0.43) | 1.37(0.31)              | 0.94                 |
| 500 MHz             | 4.38(0.51)    | 6.25(1.92)        | 91(25)   | 63(32)     | 3.33(0.55) | 1.43(0.41)              | 0.91                 |
| 600 MHz             | 4.46(0.15)    | 6.03(0.43)        | 98(9)    | 60(13)     | 3.34(0.14) | 1.35(0.09)              | 0.76                 |
| 700 MHz             | 4.45(0.12)    | 6.23(0.38)        | 104(7)   | 54(13)     | 3.30(0.12) | 1.40(0.08)              | 1.53                 |
| 800 MHz             | 4.54(0.17)    | 6.02(0.48)        | 103(13)  | 53(21)     | 3.31(0.15) | 1.33(0.09)              | 1.12                 |
| <u>1P7F.pdb:</u>    |               |                   |          |            |            |                         |                      |
| 400 MHz             | 4.48(0.38)    | 6.15(1.40)        | 103(20)  | 70(22)     | 3.31(0.42) | 1.37(0.30)              | 0.94                 |
| 500 MHz             | 4.38(0.51)    | 6.26(1.96)        | 90(23)   | 70(31)     | 3.33(0.56) | 1.43(0.42)              | 0.90                 |
| 600 MHz             | 4.46(0.15)    | 6.04(0.44)        | 96(9)    | 66(13)     | 3.34(0.14) | 1.36(0.09)              | 0.73                 |
| 700 MHz             | 4.44(0.13)    | 6.24(0.37)        | 100(7)   | 60(14)     | 3.31(0.12) | 1.41(0.07)              | 1.58                 |
| 800 MHz             | 4.54(0.17)    | 6.02(0.49)        | 100(13)  | 59(20)     | 3.31(0.15) | 1.33(0.10)              | 1.12                 |

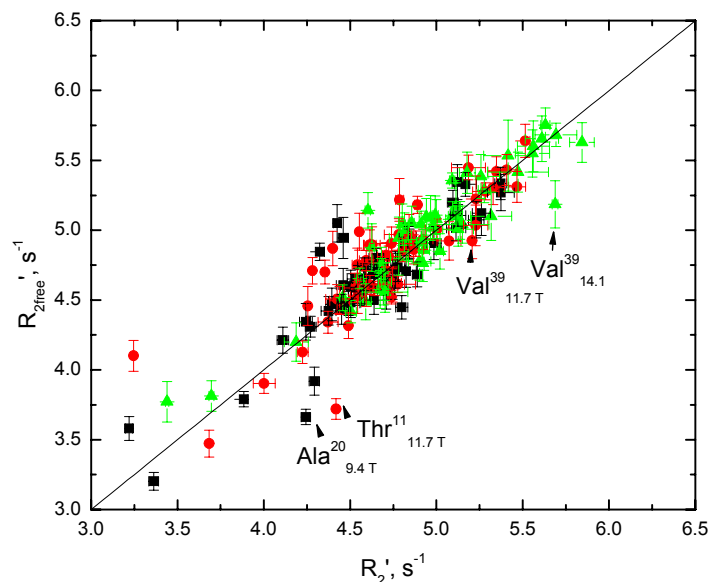
The good agreement (within the experimental errors) between the diffusion tensors determined at different fields indicates that there is no significant difference in the experimental conditions (in particular, temperature) between the measurements on different spectrometers—this then justifies the simultaneous analysis of these relaxation data acquired at various fields for the purpose of extracting field-independent parameters, such as the CSA,  $S^2$  etc. Note also that there is practically no difference between the diffusion tensors derived using the crystal and RDC-refined

crystal structure of GB3 (Table 7.1.1a,b). Also there is no significant difference between the diffusion tensor obtained from a simultaneous (global) fit of all the data and the result of averaging the diffusion tensors obtained at each field (Table 7.1.1a). Therefore for our LS analyses, we used the diffusion tensor resulting from the simultaneous fit of all data.

#### **7.1.4 $R'_{2\text{free}}$ versus $R_2$ ' as Evidence there is no Significant Conformational Exchange Contribution to $R_2$ in Measurements at Five Fields.**

A conformational exchange contribution to  $R_2$  in the case of fast exchange has the same field dependence as the  $\Delta\sigma^2$  term (e.g. Eqs.7.2.7 and 7.2.10 below), and special care is required in order to separate them. We assume throughout this paper that the conformational exchange contribution to  $R_2$  is negligible, which (as discussed in Chapter 5 based on data at 600 MHz) holds for all residues in GB3 except possibly Val39. As shown in Figure 7.1.1, this result is supported by relaxation rates measured at 400 MHz, 500 MHz.

**Figure 7.1.1 The agreement between the measured  $R_2$ s and their reconstructed “exchange-free” values,  $R_{2\text{free}} = R_1' \cdot \eta_{xy} / \eta_z$ .** Shown is the agreement between the values of  $R_2'$  and  $R_{2\text{free}}$  in GB3 at 9.4 (black squares), 11.7 T (red circles) and 14.1 T (green triangles). Here  $R_{\text{ex}}$  motions are manifest as shifts of data points to the right of the diagonal and in principle, should be proportional to the strength of the applied magnetic field squared. The only residues which show such shifts in GB3 are indicated with arrows and text labels. Ala20 shows a large shift to the right at 9.4 T, but no such shift at higher fields, indicating that this shift is most likely due to experimental error rather than  $R_{\text{ex}}$  motion. Similarly for the large shift of Thr11 at 11.7 T. Only Val39 in GB3 has a shift that gets increasingly larger with field strength.



## **7.2 Model independent methods for measurement of $^{15}\text{N}$ CSA from Relaxation Rates at Several Field Strengths.**

We derived the  $^{15}\text{N}$  chemical shielding anisotropies from measured relaxation and cross-correlation rates using three different model-of-motion independent methods outlined below. These methods for determining the  $^{15}\text{N}$  CSA involve linear fits of combinations of relaxation data at several field strengths versus the square  $^{15}\text{N}$  Larmor frequency or the  $^1\text{H}$  Larmor frequency and do not involve any assumptions about the type of motion of the NH bond. Still there are several assumptions involved in these methods, the validity of which affects the accuracy of the derived CSAs. These assumptions are discussed in detail in section 7.6.1.

### 7.2.1 The R/η method.

This method is a generalization of that of<sup>51</sup> and is based on the idea that the ratio of the corresponding cross-correlation and auto-relaxation rates is independent, to a good approximation, of the spectral densities  $J(\omega)$ <sup>50,101</sup>:

$$\frac{R_2'}{\eta_{xy}} = \frac{R_1'}{\eta_z} = \frac{d^2 + c^2}{2dc_g} \quad (7.2.1)$$

Here  $c_g = -\omega_N \cdot \Delta\sigma_g/3$  and represents the <sup>15</sup>N CSA contributions to cross-correlation rates, where<sup>74</sup>

$$|\Delta\sigma| = [\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 - (\sigma_{xx}\sigma_{yy} + \sigma_{xx}\sigma_{zz} + \sigma_{yy}\sigma_{zz})]^{1/2}, \quad (7.2.2)$$

is related to  $c$  as defined in Chapter 2, and:

$$\Delta\sigma_g = (\sigma_{zz} - \sigma_{yy})P_2(\cos\beta_z) + (\sigma_{xx} - \sigma_{yy})P_2(\cos\beta_x); \quad (7.2.3)$$

$\sigma_{ii}$  are the principal values of the <sup>15</sup>N CST.  $\Delta\sigma_g$  has the meaning of a “projection” of the CSA tensor onto the NH vector and can be represented as  $\Delta\sigma$  times an orientation factor. Under the assumption of an axial symmetry of the <sup>15</sup>N CST ( $\sigma_{xx}=\sigma_{yy}=\sigma_{\perp}$ ,  $\sigma_{zz}=\sigma_{\parallel}$ ), Eqs.7.2.2 and 7.2.3 simplify into their more conventional forms (e.g.<sup>50</sup>):

$$\Delta\sigma = \sigma_{\parallel} - \sigma_{\perp} \text{ and } \Delta\sigma_g = \Delta\sigma \cdot P_2(\cos\beta_z). \quad (7.2.4)$$

The primes in Eq.7.2.1 indicate “reduced” relaxation rates (Eqs. 3.5.17-3.5.18). Equation (7.2.1) can be recast to yield a linear dependence on  $\omega_N^2$ ,

$$\frac{2\omega_N R_2'}{\eta_{xy}} = \frac{2\omega_N R_1'}{\eta_z} = -\frac{3d}{\Delta\sigma_g} - \frac{(\Delta\sigma)^2}{3d \cdot \Delta\sigma_g} \omega_N^2. \quad (7.2.7)$$

which can then be fit to a straight line,  $m \cdot x + b$  (where  $x = \omega_N^2$ ), using a simple linear regression. This form allows a direct determination of  $\Delta\sigma_g$  and  $\Delta\sigma$  from the intercept  $b$  and the slope  $m$  of this line:

$$\Delta\sigma_g = -3d / b; \quad (7.2.8)$$

$$\Delta\sigma = -3d (m / b)^{1/2}. \quad (7.2.9)$$

The choice of the sign in Eq.7.2.9 reflects negative  $^{15}\text{N}$  CSA, in agreement with solid-state NMR data. For an axially symmetric  $^{15}\text{N}$  CST this gives <sup>51,52</sup> (cf. Eq.7.2.4)  $\sigma_{\parallel} - \sigma_{\perp} = -3d (m/b)^{1/2}$  and  $P_2(\cos\beta_z) = (m \cdot b)^{-1/2}$ .

### 7.2.2 The $2R_2 - R_1$ method.

This method is based on a quadratic field dependence of the following combination of the auto-relaxation rates (e.g. <sup>52</sup>),

$$2R_2' - R_1' = 4 d^2 J(0) + (4/9) J(0) (\Delta\sigma)^2 \omega_N^2, \quad (7.2.10)$$

which allows determination of  $J(0)$  and  $\Delta\sigma$  from the slope  $m$  and the intercept  $b$  of the line  $m \cdot \omega_N^2 + b$  representing a linear dependence of  $2R_2' - R_1'$  on  $\omega_N^2$ :

$$J(0) = b / (4d^2); \quad (7.2.11)$$

$$\Delta\sigma = -3d (m / b)^{1/2}. \quad (7.2.12)$$

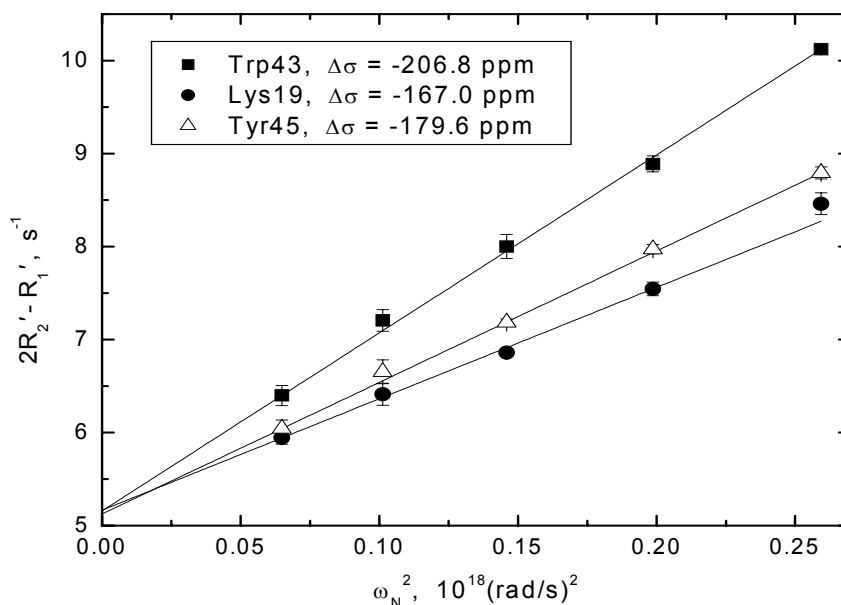
In this method, the spectral density  $J(0)$  is determined from the intercept of the fitting line, and therefore is independent of the  $^{15}\text{N}$  CSA. Given the results of Chapter 6, we assume throughout this analysis that the conformational exchange contribution to  $R_2$  is negligible, which holds for all residues in GB3 except possibly Val39 <sup>152</sup>. When present, conformational exchange contribution (in the case of fast exchange) has the same field dependence as the  $(\Delta\sigma)^2$  term (e.g. Eqs.7.2.7,7.2.10), and special care is



required in order to separate the two contributions<sup>52,101</sup>. An illustration of this fit for three residues in GB3 is shown in Figure 7.2.1.

**Figure 7.2.1 Representative fits of the dependence of  $2R_2' - R_1'$  on  $\omega_N^2$ .**

Shown are fits from the  $2R_2 - R_1$  method for three residues in GB3. This plot also illustrates the variation in the  $^{15}\text{N}$  CSA values between these residues. The amides shown here have very similar values of  $J(0)$ , as evidenced by the fact that they have the same intercept  $b$  (cf. Eq. 7.2.11), but exhibit strikingly different slopes reflecting the difference in their CSA values (Eq. 7.2.12). The plots of  $2R_2' - R_1'$  versus  $\omega_N^2$  for all residues in GB3 can be found in the Appendix. The error bars here and in all other figures represent standard errors (corresponding to 68.3% confidence intervals).



### 7.2.3 The $2\eta_{xy} - \eta_z$ method.

This method utilizes a linear field dependence of the combination of the cross-correlation rates:

$$2\eta_{xy} - \eta_z = -(8/3) d \Delta\sigma_g \cdot J(0) \omega_N = m \cdot \omega_N, \quad (7.2.13)$$

which allows determination of the product,  $\Delta\sigma_g \cdot J(0)$ , directly from the slope  $m$  of the fitting line with zero intercept:

$$\Delta\sigma_g \cdot J(0) = -m \cdot 3/(8d). \quad (7.2.14)$$

This method has an advantage over the abovementioned methods in that it is not affected by possible conformational exchange contribution to  $R_2$  and it does not require correction for the high-frequency components of the spectral density (cf. Eqs. 7.2.5, 7.2.6). The drawback is that it does not allow separate determination of  $\Delta\sigma_g$  and  $J(0)$ . However, if one of these parameters is known (e.g.  $J(0)$  from the  $2R_2$ - $R_1$  method), then the other one (in this case,  $\Delta\sigma_g$ ) can be directly obtained from Eq. 7.2.14.

#### 7.2.4 Robust Analysis of Data.

The methods described above usually rely on a least-squares fit of experimental data. Given the small number of available experimental data points per residue, the results of such fit are susceptible to experimental errors. Measures were taken to ensure that the conditions of each experiment were identical within practical limits; however, there are outlying data points in several residues, as can be seen, for example, from the linear regression plots (Appendix A). These deviations do not seem to come from the random noise in the spectra, but rather are a result of spectral artifacts caused by baseline drift, water suppression problems etc, the distribution of which is unknown and cannot be readily determined from the small sample of measurements. Least-squares fits (including linear regression) are particularly susceptible to outliers<sup>117,153,154</sup>, as their contributions to the target function increase as a square of the deviation from the fitting curve. In light of this, for each method of deriving the CSA, in addition to the “standard” least-squares regression analysis, two so-called “robust” regression methods<sup>117,153</sup> were used to obtain alternative values of the CSA and other pertinent parameters, with slightly different weights given to

outlying data points. A least-squares regression involves minimization of the target

function,  $\rho(z) = \frac{1}{2} z^2$ , where  $z$  is given by  $z = \frac{y_i^{meas} - y^{pred}(x_i)}{\delta y_i}$ , and  $y_i^{meas}$  and  $y^{pred}(x_i)$

are the measured and predicted data, respectively, for a given residue, and  $\delta y_i$  is the experimental uncertainty in  $y_i^{meas}$ . For this type of  $\rho(z)$ , the more deviant the point from the model, the greater the weight that this point is given in the minimization.

Robust regression methods involve minimization of alternative functions of  $z$ . Here we use two such functions as the target of the minimization<sup>117,153</sup>: (1) the absolute

value of  $z$  ( $\rho(z) = |z|$ ), in which all deviant points are given the same relative weight,

and (2)  $\rho(z) = \log(1 + \frac{1}{2} z^2)$ , where the relative weight given to deviant points

initially increases with deviation (while  $z < \sqrt{2}$ ) and then decreases so that those points which are the furthest from the fitting curve are given the least relative weight.

For the majority of residues in GB3 the results of the least-squares regression and the two robust methods agreed within their estimated uncertainties. For these residues the average of the parameters from the three types of regression is reported. As the experimental uncertainties in the derived parameters we report the biggest of the errors from the least-squares fit (using standard equations<sup>117</sup> for uncertainties in linear regression parameters or Monte-Carlo simulations) and from the robust methods (using Monte-Carlo simulations), estimated by propagating the experimental errors in relaxation and cross-correlation rates.

For those few residues where the three methods disagreed (i.e. where using a different weight function for the same data set resulted in significant changes in the derived fitting parameters) no CSA is reported – except those cases where the

deviation in the CSA derived from the least-squares regression can be unambiguously ascribed to undue weight given to a single clearly outlying data point (see examples in Appendix A). For these residues, the average of the two robust methods is reported. All three fits (least-squares and the two robust methods) for each model-independent method for every amide are shown in Appendix A.

### **7.3 Analyses of Relaxation Data Using the Lipari-Szabo Approximation.**

While the methods outlined above are independent of the model of local and overall motion, the following two approaches to determination of the  $^{15}\text{N}$  CSA use a specific, so-called “model-free” or Lipari-Szabo (LS) form of the spectral density function<sup>27-29</sup> that describes the backbone dynamics in terms of an order parameter  $S$  and a correlation time  $\tau_{\text{loc}}$  of local motion (see Chapter 2). Previously, this type of inclusion of CSA in the derivation of the LS parameters has been used to assess the accuracy of overall rotational diffusion parameters<sup>151</sup>.

#### **7.3.1 “Standard” Lipari-Szabo approach (LS).**

The now standard, LS-type analysis of the relaxation data ( $R_1$ ,  $R_2$ , NOE) (see e.g.<sup>143,155</sup>) was performed using the program DYNAMICS and assuming a uniform  $^{15}\text{N}$  CSA value, as described in<sup>143,152</sup>. Up to eight motional models (listed in<sup>143</sup>) were considered per residue, depending on the number of available observables. The overall tumbling of GB3 was assumed anisotropic, described by the average diffusion tensor shown in Table 7.1.1. For amides in the loop regions, where the NH vector orientation is probably less well defined than in the elements of secondary structure we assumed the overall tumbling was assumed isotropic, in order to avoid bias by a

particular loop conformation captured in the crystal structure. Using the anisotropic diffusion model and crystal structure coordinates for residues in the loop regions resulted in slightly different values of the order parameters<sup>152</sup> but did not alter the conclusions of the analysis. The same approach was also adopted for the other LS-based models throughout this paper.

### **7.3.2 Lipari-Szabo Approach Including the CSA (LS-CSA).**

This approach is an extension of the “standard” LS analysis of the relaxation data ( $R_1$ ,  $R_2$ , NOE) (see above) that here includes site-specific  $^{15}\text{N}$  CSA ( $\Delta\sigma$ ) as an additional adjustable parameter. The LS-CSA method, therefore, yields  $\Delta\sigma$  and the conventional LS parameters (e.g.  $S^2$ ,  $\tau_{\text{loc}}$ ) and possibly  $R_{\text{ex}}$ , depending on the model selection for local dynamics. Up to eight motional models (listed in<sup>143</sup>) were considered per residue, depending on the number of available observables. For these purposes, the recent version of our computer program DYNAMICS<sup>152</sup> that already accounts for the overall rotational anisotropy was upgraded to include  $\Delta\sigma$  as an additional fitting parameter in a simplex-based optimization.

The robustness of this procedure of deriving the  $\Delta\sigma$  was tested on 1,000 sets (per model) of synthetic relaxation data ( $R_1$ ,  $R_2$ , NOE at the five field strengths) containing 1% experimental noise. The range of the input parameters for simple LS models was:  $S^2$  from 0.6 to 1,  $\tau_{\text{loc}}$  from 0 to 100 ps (typical range of values for elements of secondary structure), and  $\Delta\sigma$  from -100 to -300 ppm. The output order parameters and the  $\Delta\sigma$  were within 4.38% (mean 0.004%, std 1.08%) and 6.68% (mean -0.012%, std 1.71%), respectively, from their input values, though only 94.9%

of the data could be fit to within a 95% confidence level with this level of noise. In the case of the “extended model-free” model <sup>29</sup>, the fast dynamics were characterized by  $S_{\text{fast}}^2$  from 0.7 to 1 (with  $S^2 = S_{\text{slow}}^2 \cdot S_{\text{fast}}^2 < S_{\text{fast}}^2$ ) and  $\tau_{\text{fast}}$  from 0 to 100 ps, while the slow motions had  $S_{\text{slow}}^2$  from 0.6 to 1 and  $\tau_{\text{slow}}$  from 200 to 500 ps. Here the output order parameters and the  $\Delta\sigma$  were within 4.78% (mean 0.008%, std 1.11%) and 8.93% (mean -0.02%, std 1.86%), respectively, from their input values, and 95.9% of the data could be fit to within a 95% confidence level with this level of noise. No  $R_{\text{ex}}$  contributions to  $R_2$  were included in the simulation. From these analyses, we concluded that the order parameter and CSA could be fit to within reasonable uncertainty with the existing errors in the experimental relaxation data.

### 7.3.3 Lipari-Szabo Analysis of Spectral Densities (LS-SDF).

The CSA values were also derived by simultaneous fitting of the spectral densities measured at all five fields to a LS spectral density,  $J_{\text{LS}}(\omega)$  <sup>27</sup>, that describes local dynamics in terms of  $S^2$  and  $\tau_{\text{loc}}$  only. The  $J_{\text{LS}}(\omega)$  values included the effect of the overall rotational anisotropy <sup>70,156</sup>, calculated from the diffusion tensor characteristics (Table 7.1.1) and the orientation of a given NH vector reconstructed according to the crystal structure of GB3 (1IGD.pdb). For each residue, the experimental values of the spectral density function  $J(\omega)$  at  $\omega=0$ ,  $\omega_{\text{N}}$ , and  $0.87\omega_{\text{H}}$  were directly derived from the relaxation data ( $R_1$ ,  $R_2$ , NOE) at each field strength using the reduced spectral density approximation <sup>78,79</sup>, as outlined in Chapter 2 (Eqs. 2.5.16-2.5.18). Altogether this resulted in 15 values of  $J(\omega)$  per residue, five of which were  $J(0)$  values derived from different-field measurements and which are expected

to be the same within experimental precision. The LS parameters, ( $S^2$ ,  $\tau_{\text{loc}}$ ), and the  $^{15}\text{N}$  CSA value for each residue were obtained from an unconstrained nonlinear minimization of the following target function:

$$\chi_{MF}^2 = \sum_i \left[ \frac{J(\omega_i) - J_{LS}(\omega_i)}{\delta J_i} \right]^2 \quad (7.3.4)$$

where the sum is over all available  $\omega_i$  values for a given residue, and  $\delta J_i$  represents the experimental error in  $J(\omega_i)$ . This method is analogous to the “classical” LS analysis except that reduced spectral densities are being used and the CSA is an additional fitting parameter.

#### **7.4 Site-Specific CSAs in GB3.**

##### **7.4.1 Site Specific $\Delta\sigma$ , $\Delta\sigma_g$ , and the Principal Values and Orientations of the Assumed Axially Symmetric Chemical Shielding Tensors.**

The values of the residue-specific  $^{15}\text{N}$  CSAs measured in GB3 using the model-independent methods ( $2R_2 - R_1$  and  $R/\eta$ ) are shown in Figure 7.4.1a versus residue number. The secondary structure of the protein is indicated by bars at the top of the figure.

**The  $2R_2 - R_1$  Method.** The  $^{15}\text{N}$   $\Delta\sigma$  values and the spectral density  $J(0)$  were determined directly from the observed field dependence of the combination of reduced auto-relaxation rates,  $2R_2' - R_1'$  (as in Fig. 7.2.1). Relaxation data ( $R_1$ ,  $R_2$ , and NOE) at all five fields were used for each residue. The data were fitted to a linear dependence on  $\omega_N^2$  (Eq.10) using the three linear regression methods (least-squares and two robust methods) as discussed earlier; the quality of the fit for each residue is

shown in Appendix A. All three regression methods had good agreement (both slope and intercept agreed within the experimental uncertainty) for 38 out of 50 residues in GB3. For an additional 9 residues (Leu12, Ala20, Val21, Gly38, Asp40, Asp47, Ala48, Thr49, and the C-terminal residue, Glu56) the two robust methods agreed within their experimental uncertainties (68.3% confidence interval). Only for 3 residues (Lys10, Gly41, and Lys50, all of which are in the loops in GB3) can no definitive CSA be reported because all three regression methods disagree for the  $2R_2' - R_1'$  fit.

The average site-specific  $^{15}\text{N}$  CSA values from the three fits are presented in Fig.7.4.1 (solid squares), the values of  $J(0)$  are shown later on in Fig.7.5.2 (solid squares) in comparison with  $J(0)$  values calculated from order parameters. The site-specific  $^{15}\text{N}$  CSAs from this method range from  $-111.3 \pm 1.7$  ppm (Leu12) to  $-241.0 \pm 8.7$  ppm (Phe52), with a mean of  $\langle \Delta\sigma \rangle = -174.2$  ppm and a standard deviation of 22.2 ppm. The median  $\Delta\sigma$  is  $-175.4$  ppm, in good agreement with the mean, indicating that the mean is not dominated by a small number of outliers (Table 7.4.1). The average estimated relative uncertainty is 2.67% for  $J(0)$  and 3.44% (or 6.0 ppm) for  $\Delta\sigma$ .

**The  $R/\eta$  Method.** This method is based on the field dependence of the ratio of the (reduced) auto-relaxation rate ( $R_2'$  or  $R_1'$ ) and the corresponding  $^{15}\text{N}$  CSA/dipolar cross-correlation rate ( $\eta_{xy}$  or  $\eta_z$ , respectively), Eqs.1-9. Both  $R_2'/\eta_{xy}$  and  $R_1'/\eta_z$  ratios are expected to have the same values (Eq.7.2.1), therefore these data were analyzed together. The analysis included  $R_2'/\eta_{xy}$  data at four fields and  $R_1'/\eta_z$  at three fields for each residue. Using both  $R_2'/\eta_{xy}$  and  $R_1'/\eta_z$  data improves the accuracy of analysis by



increasing the number of data points included in the fit. In addition, the  $R_1'/\eta_z$  values have the advantage of being free of any contribution from conformational exchange. The quality of the fit for each residue in GB3 is shown in Appendix A. All three regression methods had good agreement (both slope and intercept agreed within the experimental uncertainty) for 37 out of 50 amides in GB3. For an additional 7 residues (Gly9, Thr11, Lys13, Ala26, Gly38, Phe52, and the C-terminal Glu56) the two robust methods agreed within their experimental uncertainties (68.3% confidence interval). For 6 residues (Leu12, Ala20, Asp40, Gly41, Ala48, and Thr49, all of which are in loop/turn regions of GB3), no CSA is reported here because all three regression methods disagree in the  $R/\eta$  fit. The  $^{15}\text{N}$  CSA values ( $\Delta\sigma$ ) obtained using this approach are shown in Fig. 7.4.1, the values of  $\Delta\sigma_g$  are presented in Fig. 7.4.2. These  $^{15}\text{N}$  CSAs range from  $-127.9 \pm 4.0$  ppm (Gly38) to  $-237.9 \pm 11.1$  ppm (Phe52), with a mean value of  $-177.4$  ppm and a standard deviation of 19.5 ppm. The median is  $-178.4$  ppm. The average estimated level of the experimental errors is 4.23% (or 7.5 ppm) for  $\Delta\sigma$ .

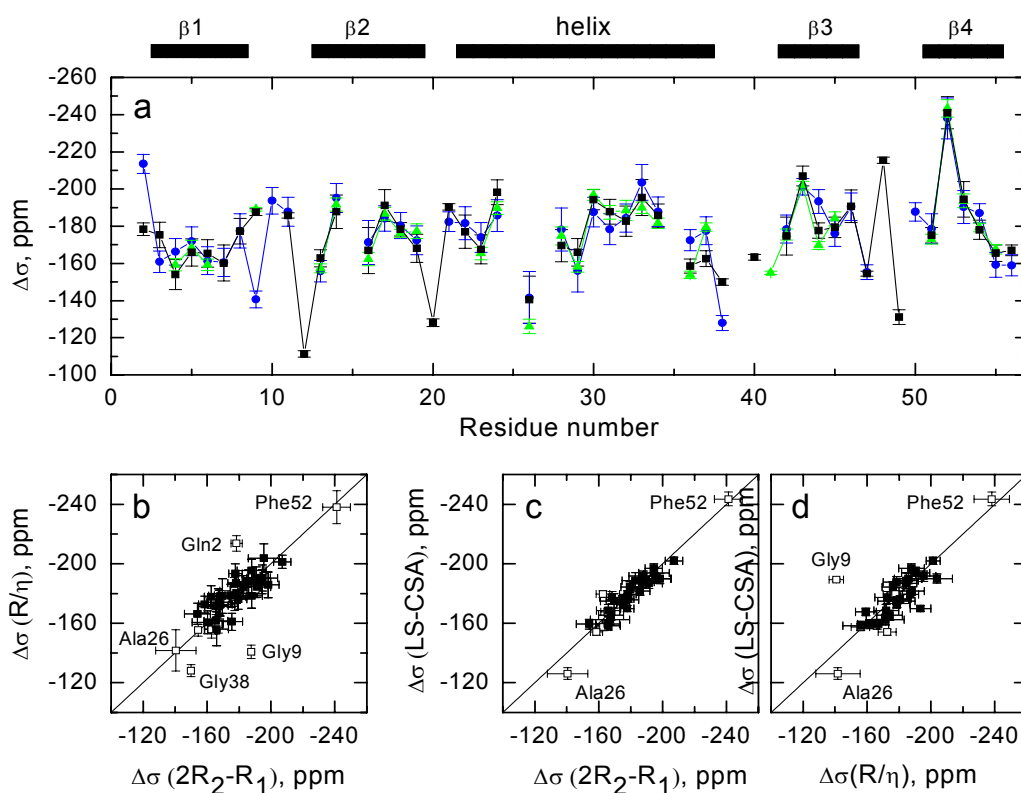
The angles  $\beta_z$  derived from these  $\Delta\sigma$  and  $\Delta\sigma_g$  values assuming axial symmetry of the  $^{15}\text{N}$  CST are shown in Fig. 7.4.2c (black squares). The range of  $\beta_z$  values is from  $7.5^\circ$  (Val6) to  $27.6^\circ$  (Thr11) with a mean value of  $19.9^\circ$  and standard deviation of  $4.5^\circ$ , in agreement with the  $\beta_z$  values observed in ubiquitin. Very similar  $\beta_z$  values were also determined from a combination of the  $\Delta\sigma_g$  values from the  $2\eta_{xy}-2\eta_z$  method with the  $\Delta\sigma$  values from  $2R_2-R_1$  (see below).

Note that using the mean of  $R_2'/\eta_{xy}$  and  $R_1'/\eta_z$  as the  $R/\eta$  value at a given field (where both data are available at 9.1, 11.7, and 14.1 T) resulted in the CSA values

from  $-127.9$  to  $-237.9$  ppm with a mean CSA of  $-177.4$  ppm and a standard deviation of  $19.5$  ppm. These results have an overall correlation coefficient of  $0.97$  to CSA values obtained using the individual measurements. Fitting the  $R_2'/\eta_{xy}$  values alone gave  $^{15}\text{N}$  CSA values in the range from  $-140.5$  to  $-234.8$  ppm, with a mean of  $-179.2$  ppm and a standard deviation of  $19.2$  ppm, with a correlation coefficient of  $0.91$  to the CSAs derived from both transverse and longitudinal data. The  $R_1'/\eta_z$  data alone yielded CSAs from  $-129.9$  to  $-251.6$  ppm, with a larger absolute values of the mean ( $-185.5$  ppm) and standard deviation ( $23.9$  ppm). These data show a poor correlation ( $r=0.13$ ) with the CSAs obtained from both transverse and longitudinal data together, which likely reflects a lesser accuracy of the  $R_1/\eta_z$  data alone due to a narrower range of magnetic fields covered by the  $\eta_z$  measurements.

The  $^{15}\text{N}$  CSAs obtained by the  $R/\eta$  method are expected to be independent of the magnitude of the spectral density function. Indeed, the Pearson's correlation coefficient  $r$  between the  $J(0)$  values derived from the  $2R_2-R_1$  method (these values are independent of  $\Delta\sigma$ ) and the CSA values from the  $R/\eta$  approach was  $-0.23$ .

**Figure 7.4.1 (a) The site-specific  $^{15}\text{N}$  CSAs, from the  $2\text{R}_2\text{-R}_1$  method (black squares),  $\text{R}/\eta$  method (blue circles), and the LS-CSA method (green triangles) versus residue number.** The secondary structure of GB3 is indicated at the top of the panel. (b) Correlation between  $^{15}\text{N}$  CSA values measured using the model-independent methods,  $2\text{R}_2\text{-R}_1$  and  $\text{R}/\eta$ . The Pearson's correlation coefficient  $r$  for these two data sets is 0.79; 81% of these CSA data agree within the experimental uncertainties. These values improve to  $r=0.80$  and 87% agreement if only those data (shown as solid squares) where the least-squares fits pass the 95%-confidence level  $\chi^2/\text{df}$  cutoff are considered. (c) Correlation between the CSAs from  $2\text{R}_2\text{-R}_1$  and LS-CSA methods. The correlation coefficient is 0.95; it decreases to  $r=0.93$  if only those fits that pass the  $\chi^2/\text{df}$  cutoff (solid squares) are included, though the percent agreement improves from 94% to 96%. (d) Correlation between the results from  $\text{R}/\eta$  and LS-CSA methods. The correlation coefficient is 0.80 and remains unchanged when the  $\chi^2/\text{df}$  cutoff is applied (solid squares). The percent agreement increases from 84% for all considered residues to 88% for those residues with the  $\chi^2/\text{df}$  below the cutoff value. In all correlation plots (panels b-d) the solid symbols represent values obtained for least squares fits that passed the  $\chi^2/\text{df}$  cutoff while open symbols correspond to the remaining residues. Outliers and extreme values of the CSA are labeled. Note that those few residues that show significant differences in the CSA values between the methods are all located in the loops/termini. Also in the loops are all residues where only one out of the three methods resulted in an acceptable fit (panel a).

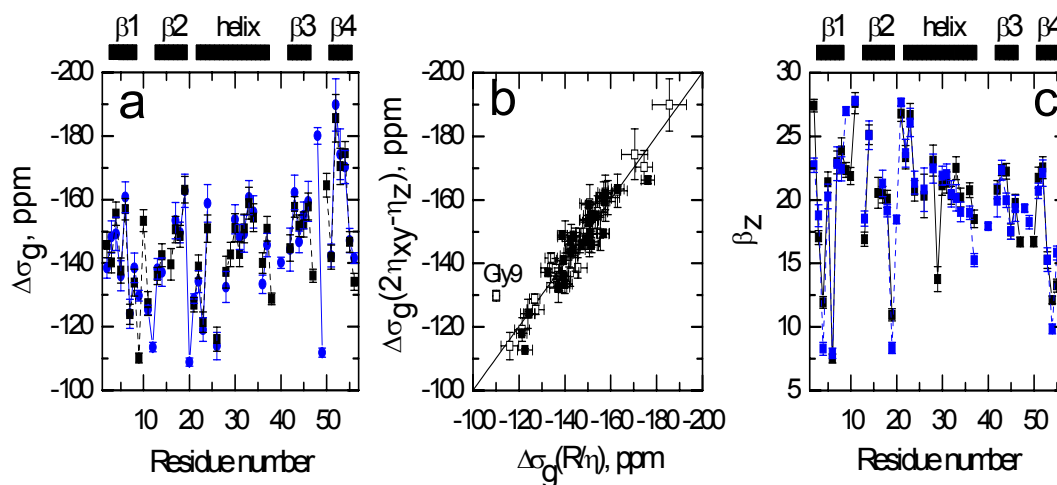


**Quality Control Using the  $2\eta_{xy}\text{-}\eta_z$  Method.** The field dependence of the cross-correlation data alone yields the product of  $\Delta\sigma_g$  and  $J(0)$ . This analysis is independent of the auto-relaxation data. We then used the value of  $J(0)$  derived from the  $2\text{R}_2\text{-R}_1$  method (this value is independent of  $\Delta\sigma$ ) to obtain  $\Delta\sigma_g$  (Fig.7.4.2). The  $\Delta\sigma_g$  values

thus obtained range from  $-107.2 (\pm 1.2)$  ppm for Leu12 to  $-186.1 (\pm 1.0)$  ppm for Ala34, with the mean value of  $-154.4$  ppm, and a median at  $-154.1$  ppm. These values were then compared with the  $\Delta\sigma_g$  values derived from the  $R/\eta$  approach, which are independent of  $J(0)$ . The excellent agreement ( $r=0.94$  for all residues, and  $0.96$  for filled symbols, Fig.7.4.2b) between the values of the same parameter determined independently from different sets of measurements is a quality control measurement. Assuming axial symmetry of the  $^{15}\text{N}$  CST, and using  $\Delta\sigma$  values from the  $2R_2-R_1$  method, we determined the angle  $\beta_z$  between the unique (least shielded) component of the tensor and the NH bond vector (Fig.7.4.2c). These  $\beta_z$  values are in very good agreement ( $r=0.93$ ) with  $\beta_z$  derived from the  $R/\eta$  method described above.

**Figure 7.4.2 The values of  $\Delta\sigma_g$  and the  $\beta_z$  angles from the  $R/\eta$  and  $2\eta_{xy}-\eta_z$  methods**

(a) Measured site-specific  $^{15}\text{N}$   $\Delta\sigma_g$  values for GB3 from the  $R/\eta$  (squares) and the  $2\eta_{xy}-\eta_z$  methods (circles). The  $\Delta\sigma_g$  values range from  $-108.9$  ppm (Ala20,  $2\eta_{xy}-\eta_z$ ) to  $-189.8$  ppm (Phe52,  $2\eta_{xy}-\eta_z$ ). (b) Correlation between  $\Delta\sigma_g$  values measured using the  $R/\eta$  and  $2\eta_{xy}-\eta_z$  methods. The correlation coefficient is  $0.94$  for all residues and  $0.95$  for only those fits that pass the  $\chi^2/\text{df}$  cutoff. (c)  $\beta_z$  angles (in degrees) determined from the  $R/\eta$  method (squares) and by combining the  $\Delta\sigma_g$  values from the  $2\eta_{xy}-\eta_z$  method with the  $\Delta\sigma$  values from  $2R_2-R_1$  (circles). The Pearson's correlation coefficient for the agreement of the  $\beta$  angles from these two measurements is  $0.94$ . The derivation of  $\beta_z$  assumed axial symmetry of the  $^{15}\text{N}$  chemical shielding tensor. The secondary structure of GB3 is indicated on the top of panels (a) and (c).



**Methods Based on the Lipari-Szabo Form of the Spectral Density Function.** There is good agreement between the results of the LS-CSA and LS-SDF methods: for the residues in the secondary structure, the CSAs from the two methods agree within their errors and have a correlation coefficient of 0.98. The order parameters and  $\tau_{\text{loc}}$  values derived using these methods agree within their respective errors for all but two residues (Ala23 and Lys28) in the secondary structure. For those residues where there is good agreement, this indicates that the use of approximate reduced spectral densities does not significantly alter the values of the fit parameters.

Furthermore, the CSA values from these two approaches based on the LS form of the spectral density function are in good agreement with the results of the model-independent approaches (Fig. 7.4.1c,d). For all residues in GB3, the Pearson's correlation coefficient is 0.95 between the CSAs from the LS analyses and the  $2R_2-R_1$  method and 0.80 between the CSA values from the LS analyses and those measured using the  $R/\eta$  method. The range of  $^{15}\text{N}$  CSAs obtained from all abovementioned methods for each residue in GB3 is shown in Fig. 7.4.3, together with a histogram of the average CSA values (from the three determination methods) for each residue. These site-specific  $^{15}\text{N}$  CSA values were then combined with the isotropic chemical shift data in order to reconstruct the individual components of the  $^{15}\text{N}$  CST in GB3, assuming axial symmetry of the tensor (Tables 7.4.2a,b).

The lower correlation between the LS analysis and the  $R/\eta$  methods in comparison with that between the LS and the  $2R_2-R_1$  method is potentially interesting. This might be due to a difference in the motions or chemical shifts sampled by cross correlation and auto-relaxation rates. Another possible explanation

of this is that the LS and  $2R_2-R_1$  analyses use the same experimental data ( $R_1$ ,  $R_2$ , and the NOE at five field strengths), while the  $R/\eta$  analysis additionally includes the cross correlation rates. The good correlation between the LS and  $2R_2-R_1$  results implies that the difference in the type of analyses does not make a significant difference in the derived values of the CSAs, whereas the somewhat low correlation between the LS and  $R/\eta$  CSAs could be due to the different types of analyses or to different sampling of motions by the different measured rates (CCRs versus auto-relaxation rates), and it is difficult to discriminate between these two possibilities from this data.

**Table 7.4.1. Statistics of the  $^{15}\text{N}$  CSA values in GB3 determined here using several methods.**

| Method                          | Analyzed set of residues                        | Number of residues | $\max(\Delta\sigma)^a$ (ppm) | $\min(\Delta\sigma)^b$ (ppm) | $\langle\Delta\sigma\rangle^c$ (ppm) | $\mu^d$ (ppm) | median <sup>e</sup> (ppm) | $\langle\delta\Delta\sigma\rangle^f$ (ppm) | std( $\Delta\sigma$ ) <sup>g</sup> (ppm) | $\Lambda^h$ (ppm) |
|---------------------------------|---|--------------------|------------------------------|------------------------------|--------------------------------------|---------------|---------------------------|--|--|-------------------|
| 2R <sub>2</sub> -R <sub>1</sub> | All <sup>i</sup>                                | 47                 | -111.3                       | -241.0                       | -174.2                               | -173.9        | -175.4                    | 6.0  | 22.2                                     | 21.4              |
|                                 | $\chi^2/\text{df}_{\text{fit}} < 95\%$ cutoff   | 32                 | -154.0                       | -207.0                       | -178.1                               | -178.2        | -178.9                    | 7.0  | 12.9                                     | 10.6              |
|                                 | $\alpha$ -helix <sup>j</sup>                    | 11                 | -140.4                       | -198.2                       | -175.8                               | -176.4        | -177.0                    | 7.6  | 18.1                                     | 14.1              |
|                                 | $\beta$ -strands <sup>j</sup>                   | 19                 | -154.0                       | -241.0                       | -180.3                               | -180.2        | -177.5                    | 7.3  | 19.1                                     | 16.3              |
| R/ $\eta$                       | All <sup>i</sup>                                | 44                 | -127.9                       | -237.9                       | -177.4                               | -177.2        | -178.4                    | 7.5  | 19.5                                     | 17.6              |
|                                 | $\chi^2/\text{df}_{\text{fit}} < 95\%$ cutoff   | 33                 | -155.7                       | -203.5                       | -178.2                               | -178.2        | -178.3                    | 7.8  | 12.5                                     | 10.2              |
|                                 | $\alpha$ -helix <sup>j</sup>                    | 11                 | -141.6                       | -203.5                       | -177.6                               | -179.3        | -178.3                    | 9.2  | 16.7                                     | 8.3               |
|                                 | $\beta$ -strands <sup>j</sup>                   | 19                 | -159.2                       | -237.9                       | -181.1                               | -180.7        | -178.5                    | 7.5  | 18.3                                     | 14.7              |
| LS-CSA                          | All <sup>i</sup>                                | 32                 | -126.0                       | -243.4                       | -176.9                               | -176.9        | -176.8                    | 3.1  | 20.0                                     | 19.2              |
|                                 | $\chi^2/\text{df}_{\text{fit}} < 95\%$ cutoff   | 25                 | -158.1                       | -201.9                       | -178.3                               | -178.3        | -177.2                    | 3.3  | 12.6                                     | 11.9              |
|                                 | $\alpha$ -helix <sup>j</sup>                    | 11                 | -126.0                       | -196.9                       | -174.3                               | -174.3        | -180.5                    | 3.4  | 21.3                                     | 19.9              |
|                                 | $\beta$ -strands <sup>j</sup>                   | 16                 | -159.3                       | -243.4                       | -180.7                               | -180.6        | -175.9                    | 3.1  | 20.9                                     | 19.6              |
| Average of all 3 methods        | All <sup>i</sup>                                | 50                 | -111.3                       | -240.8                       | -174.2                               | -173.8        | -175.9                    | 7.1  | 22.2                                     | 21.2              |
|                                 | $\chi^2/\text{df}_{\text{fit}} < \text{cutoff}$ | 35                 | -155.7                       | -203.4                       | -177.7                               | -177.2        | -178.3                    | 7.9  | 11.9                                     | 9.1               |
|                                 | $\alpha$ -helix <sup>j</sup>                    | 11                 | -136.0                       | -196.3                       | -176.0                               | -177.3        | -184.6                    | 9.2  | 18.1                                     | 12.0              |
|                                 | $\beta$ -strands <sup>j</sup>                   | 20                 | -159.8                       | -240.8                       | -180.3                               | -179.9        | -177.8                    | 8.2  | 18.6                                     | 14.5              |

<sup>a</sup> The smallest absolute value of the  $^{15}\text{N}$  CSA.

<sup>b</sup> The largest absolute value of the  $^{15}\text{N}$  CSA.

<sup>c</sup> The arithmetic mean of measured values of the  $^{15}\text{N}$  CSA.

<sup>d</sup> The value of  $\mu$  that maximizes the likelihood function  $p(\mu, \Lambda)$  (Eq. 6.4.1);  $\mu$  is an estimate of the true mean of the CSA distribution.

<sup>e</sup> Median of measured values of the  $^{15}\text{N}$  CSA.

<sup>f</sup> The arithmetic mean of experimental uncertainties in the  $^{15}\text{N}$  CSA.

<sup>g</sup> The standard deviation of the measured values of the  $^{15}\text{N}$  CSA.

<sup>h</sup> The value of  $\Lambda$  that maximizes the likelihood function  $p(\mu, \Lambda)$ ;  $\Lambda$  is an estimate of the true site-to-site variability in the CSA distribution.

<sup>i</sup> All residues with acceptable agreement of regression methods (out of 50 analyzable residues, see text).

<sup>j</sup> The  $\alpha$ -helix in GB3 extends from Ala23 to Asp36 with Thr25, Glu27, and Asn35 impossible to resolve in the spectra due to overlap (hence 11 analyzable residues). The  $\beta$ -strands comprise Tyr3-Ile7, Gly14-Lys19, Val42-Asp46, and Thr51-Thr55, with Glu15 excluded due to overlap (altogether 20 analyzable residues). Gln2 was excluded from the LS analyses.

**Table 7.4.2a Site-specific components of the chemical shift tensor (assumed axially symmetric).** The error in  $\delta_{iso}$  assumed to be negligible relative to the error in  $\Delta\delta$ . Here:

$$\delta_{\parallel} = \delta_{iso} + 2\Delta\delta/3 \text{ and } \delta_{\perp} = \delta_{iso} - \Delta\delta/3, \text{ so that } \Delta\delta = \delta_{\parallel} - \delta_{\perp} \text{ and } \delta_{iso} = (2\delta_{\perp} + \delta_{\parallel})/3.$$

The experimental uncertainty indicated in parentheses.

| Residue | $\Delta\delta$ | $\delta_{iso}$ | $\delta_{\parallel}$ | $\delta_{\perp}$ |
|---------|----------------|----------------|----------------------|------------------|
| 2       | 195.94 (5.23)  | 123.13         | 253.76 (3.48)        | 57.81 (1.74)     |
| 3       | 168.11 (6.85)  | 123.87         | 235.95 (4.57)        | 67.84 (2.28)     |
| 4       | 159.82 (8.08)  | 122.34         | 228.88 (5.38)        | 69.06 (2.69)     |
| 5       | 168.7 (7.78)   | 126.29         | 238.76 (5.19)        | 70.06 (2.59)     |
| 6       | 162.06 (7.56)  | 126.87         | 234.91 (5.04)        | 72.85 (2.52)     |
| 7       | 160.36 (9.71)  | 125.32         | 232.22 (6.47)        | 71.87 (3.24)     |
| 8       | 177.36 (9.19)  | 128.83         | 247.07 (6.12)        | 69.71 (3.06)     |
| 9       | 172.46 (4.49)  | 110.05         | 225.03 (2.99)        | 52.57 (1.50)     |
| 10      | 193.68 (7.18)  | 120.44         | 249.56 (4.79)        | 55.88 (2.39)     |
| 11      | 186.81 (7.72)  | 108.68         | 233.22 (5.15)        | 46.41 (2.57)     |
| 12      | 111.26 (1.71)  | 125.43         | 199.61 (1.14)        | 88.35 (0.57)     |
| 13      | 158.6 (5.78)   | 123.36         | 229.09 (3.85)        | 70.49 (1.93)     |
| 14      | 191.5 (9.05)   | 109.06         | 236.72 (6.03)        | 45.22 (3.02)     |
| 16      | 166.83 (12.36) | 115.45         | 226.67 (8.24)        | 59.84 (4.12)     |
| 17      | 187.32 (8.52)  | 111.54         | 236.41 (5.68)        | 49.1 (2.84)      |
| 18      | 178.31 (7.02)  | 114.84         | 233.71 (4.68)        | 55.4 (2.34)      |
| 19      | 172.59 (7.72)  | 124.03         | 239.08 (5.14)        | 66.5 (2.57)      |
| 20      | 128.09 (2.04)  | 124.38         | 209.77 (1.36)        | 81.68 (0.68)     |
| 21      | 186.34 (5.33)  | 115.24         | 239.47 (3.55)        | 53.13 (1.78)     |
| 22      | 179.31 (8.97)  | 115.11         | 234.65 (5.98)        | 55.34 (2.99)     |
| 23      | 168.95 (8.03)  | 120.88         | 233.51 (5.35)        | 64.56 (2.68)     |
| 24      | 191.21 (8.63)  | 118.73         | 246.2 (5.76)         | 54.99 (2.88)     |
| 26      | 136.01 (13.94) | 125.00         | 215.67 (9.29)        | 79.66 (4.65)     |
| 28      | 174.16 (11.56) | 116.20         | 232.31 (7.71)        | 58.15 (3.85)     |
| 29      | 160.03 (11.41) | 121.97         | 228.65 (7.61)        | 68.62 (3.80)     |
| 30      | 192.95 (7.96)  | 119.51         | 248.14 (5.31)        | 55.2 (2.65)      |
| 31      | 184.55 (8.3)   | 122.67         | 245.7 (5.54)         | 61.15 (2.77)     |
| 32      | 185.17 (8.19)  | 119.39         | 242.83 (5.46)        | 57.66 (2.73)     |
| 33      | 196.28 (9.74)  | 120.19         | 251.04 (6.49)        | 54.76 (3.25)     |
| 34      | 185.03 (7.94)  | 122.28         | 245.63 (5.29)        | 60.6 (2.65)      |
| 36      | 161.59 (5.71)  | 121.00         | 228.73 (3.81)        | 67.14 (1.9)      |
| 37      | 173.19 (7.51)  | 115.09         | 230.56 (5.01)        | 57.36 (2.5)      |
| 38      | 138.95 (4.05)  | 107.95         | 200.59 (2.7)         | 61.63 (1.35)     |
| 40      | 163.32 (1.5)   | 127.72         | 236.61 (1.00)        | 73.28 (0.50)     |
| 41      | 154.88 (0.84)  | 107.04         | 210.29 (0.56)        | 55.41 (0.28)     |
| 42      | 176.23 (10.22) | 120.20         | 237.68 (6.81)        | 61.46 (3.41)     |
| 43      | 203.35 (5.34)  | 130.83         | 266.4 (3.56)         | 63.05 (1.78)     |
| 44      | 180.2 (6.41)   | 114.42         | 234.55 (4.27)        | 54.36 (2.14)     |
| 45      | 179.82 (6.83)  | 119.97         | 239.85 (4.55)        | 60.03 (2.28)     |
| 46      | 190.59 (8.77)  | 128.11         | 255.17 (5.85)        | 64.58 (2.92)     |
| 47      | 154.85 (4.01)  | 124.68         | 227.91 (2.67)        | 73.07 (1.34)     |
| 48      | 215.48 (1.77)  | 119.61         | 263.26 (1.18)        | 47.78 (0.59)     |
| 49      | 131.07 (4.04)  | 102.95         | 190.33 (2.69)        | 59.26 (1.35)     |
| 50      | 187.75 (4.98)  | 122.77         | 247.93 (3.32)        | 60.18 (1.66)     |
| 51      | 175.59 (8.09)  | 111.02         | 228.08 (5.39)        | 52.49 (2.70)     |
| 52      | 240.76 (11.05) | 130.91         | 291.41 (7.37)        | 50.66 (3.68)     |
| 53      | 192.8 (9.6)    | 117.40         | 245.93 (6.4)         | 53.13 (3.20)     |
| 54      | 181.43 (5.13)  | 122.99         | 243.95 (3.42)        | 62.51 (1.71)     |
| 55      | 164.08 (6.59)  | 123.88         | 233.26 (4.39)        | 69.18 (2.2)      |
| 56      | 162.96 (5.66)  | 133.42         | 242.07 (3.78)        | 79.1 (1.89)      |

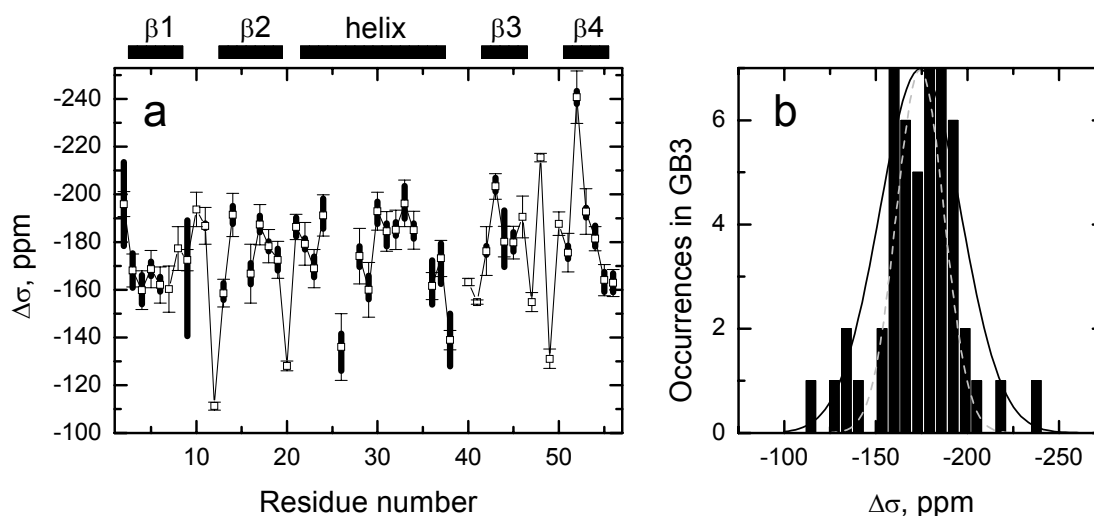


**Table 7.4.2 (b) Site-specific components of the chemical shielding tensor (assumed axially symmetric). Here:**  
 $\sigma_{\parallel} = 2\Delta\sigma/3$  and  $\sigma_{\perp} = -\Delta\sigma/3$ , so that  $\Delta\sigma = \sigma_{\parallel} - \sigma_{\perp}$  and  $2\sigma_{\perp} + \sigma_{\parallel} = 0$ .

| Residue | $\Delta\sigma$  | $\sigma_{\parallel}$ | $\sigma_{\perp}$ |
|---------|-----------------|----------------------|------------------|
| 2       | -195.94 (5.23)  | -130.63 (3.48)       | 65.31 (1.74)     |
| 3       | -168.11 (6.85)  | -112.07 (4.57)       | 56.04 (2.28)     |
| 4       | -159.82 (8.08)  | -106.55 (5.38)       | 53.27 (2.69)     |
| 5       | -168.7 (7.78)   | -112.46 (5.19)       | 56.23 (2.59)     |
| 6       | -162.06 (7.56)  | -108.04 (5.04)       | 54.02 (2.52)     |
| 7       | -160.36 (9.71)  | -106.91 (6.47)       | 53.45 (3.24)     |
| 8       | -177.36 (9.19)  | -118.24 (6.12)       | 59.12 (3.06)     |
| 9       | -172.46 (4.49)  | -114.97 (2.99)       | 57.49 (1.5)      |
| 10      | -193.68 (7.18)  | -129.12 (4.79)       | 64.56 (2.39)     |
| 11      | -186.81 (7.72)  | -124.54 (5.15)       | 62.27 (2.57)     |
| 12      | -111.26 (1.71)  | -74.17 (1.14)        | 37.09 (0.57)     |
| 13      | -158.6 (5.78)   | -105.73 (3.85)       | 52.87 (1.93)     |
| 14      | -191.5 (9.05)   | -127.66 (6.03)       | 63.83 (3.02)     |
| 16      | -166.83 (12.36) | -111.22 (8.24)       | 55.61 (4.12)     |
| 17      | -187.32 (8.52)  | -124.88 (5.68)       | 62.44 (2.84)     |
| 18      | -178.31 (7.02)  | -118.87 (4.68)       | 59.44 (2.34)     |
| 19      | -172.59 (7.72)  | -115.06 (5.14)       | 57.53 (2.57)     |
| 20      | -128.09 (2.04)  | -85.39 (1.36)        | 42.7 (0.68)      |
| 21      | -186.34 (5.33)  | -124.23 (3.55)       | 62.11 (1.78)     |
| 22      | -179.31 (8.97)  | -119.54 (5.98)       | 59.77 (2.99)     |
| 23      | -168.95 (8.03)  | -112.63 (5.35)       | 56.32 (2.68)     |
| 24      | -191.21 (8.63)  | -127.48 (5.76)       | 63.74 (2.88)     |
| 26      | -136.01 (13.94) | -90.68 (9.29)        | 45.34 (4.65)     |
| 28      | -174.16 (11.56) | -116.11 (7.71)       | 58.05 (3.85)     |
| 29      | -160.03 (11.41) | -106.69 (7.61)       | 53.34 (3.8)      |
| 30      | -192.95 (7.96)  | -128.63 (5.31)       | 64.32 (2.65)     |
| 31      | -184.55 (8.3)   | -123.03 (5.54)       | 61.52 (2.77)     |
| 32      | -185.17 (8.19)  | -123.45 (5.46)       | 61.72 (2.73)     |
| 33      | -196.28 (9.74)  | -130.86 (6.49)       | 65.43 (3.25)     |
| 34      | -185.03 (7.94)  | -123.35 (5.29)       | 61.68 (2.65)     |
| 36      | -161.59 (5.71)  | -107.73 (3.81)       | 53.86 (1.9)      |
| 37      | -173.19 (7.51)  | -115.46 (5.01)       | 57.73 (2.5)      |
| 38      | -138.95 (4.05)  | -92.63 (2.70)        | 46.32 (1.35)     |
| 40      | -163.32 (1.5)   | -108.88 (1.00)       | 54.44 (0.5)      |
| 41      | -154.88 (0.84)  | -103.25 (0.56)       | 51.63 (0.28)     |
| 42      | -176.23 (10.22) | -117.48 (6.81)       | 58.74 (3.41)     |
| 43      | -203.35 (5.34)  | -135.57 (3.56)       | 67.78 (1.78)     |
| 44      | -180.2 (6.41)   | -120.13 (4.27)       | 60.07 (2.14)     |
| 45      | -179.82 (6.83)  | -119.88 (4.55)       | 59.94 (2.28)     |
| 46      | -190.59 (8.77)  | -127.06 (5.85)       | 63.53 (2.92)     |
| 47      | -154.85 (4.01)  | -103.23 (2.67)       | 51.62 (1.34)     |
| 48      | -215.48 (1.77)  | -143.65 (1.18)       | 71.83 (0.59)     |
| 49      | -131.07 (4.04)  | -87.38 (2.69)        | 43.69 (1.35)     |
| 50      | -187.75 (4.98)  | -125.17 (3.32)       | 62.58 (1.66)     |
| 51      | -175.59 (8.09)  | -117.06 (5.39)       | 58.53 (2.7)      |
| 52      | -240.76 (11.05) | -160.5 (7.37)        | 80.25 (3.68)     |

|    |                |                |              |
|----|----------------|----------------|--------------|
| 53 | -192.8 (9.6)   | -128.53 (6.4)  | 64.27 (3.2)  |
| 54 | -181.43 (5.13) | -120.96 (3.42) | 60.48 (1.71) |
| 55 | -164.08 (6.59) | -109.38 (4.39) | 54.69 (2.2)  |
| 56 | -162.96 (5.66) | -108.64 (3.78) | 54.32 (1.89) |

**Figure 7.4.3. Site-specific  $^{15}\text{N}$  CSA values, averaged over all three methods, show significant CSA variability in GB3. (a)** Range of  $^{15}\text{N}$  CSAs for each backbone amide in GB3 from the three methods ( $2\text{R}_2\text{-R}_1$ ,  $\text{R}/\eta$ , and LS-CSA) shown as solid vertical bars. The open symbols represent the average site-specific CSA,  $\Delta\sigma$ , from the three methods; the error bars represent the maximum error from the three methods for each residue. **(b)** A histogram of the average site-specific CSA values shown in panel (a). Including these average site-specific CSA values into the analysis of the derivation of the true CSA values (Eq.7.4.1.) resulted in the true mean  $\mu = -173.8$  ppm and the site-to-site variability  $\Lambda = 21.2$  ppm (Table 7.41). The black curve represents a Gaussian distribution with the mean of  $-174.2$  ppm and the standard deviation of 22.2 ppm. The dashed curve is also a Gaussian, with the same mean but with a standard deviation of 13.0 ppm – this curve corresponds to the case when all seven outliers in panel (b) are taken out.

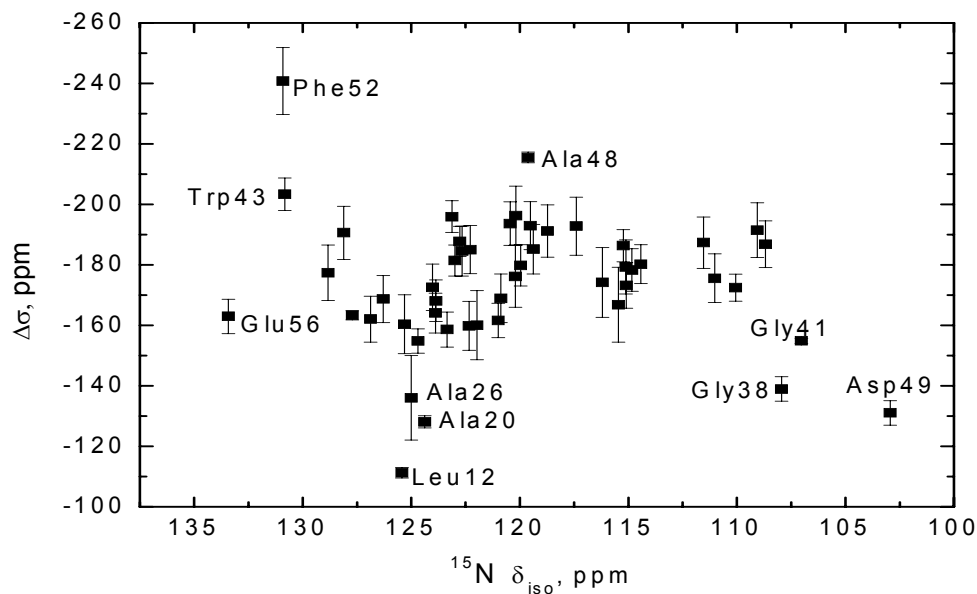


#### 7.4.2 Correlation of $\Delta\sigma$ with Isotropic Chemical Shift and Structure.

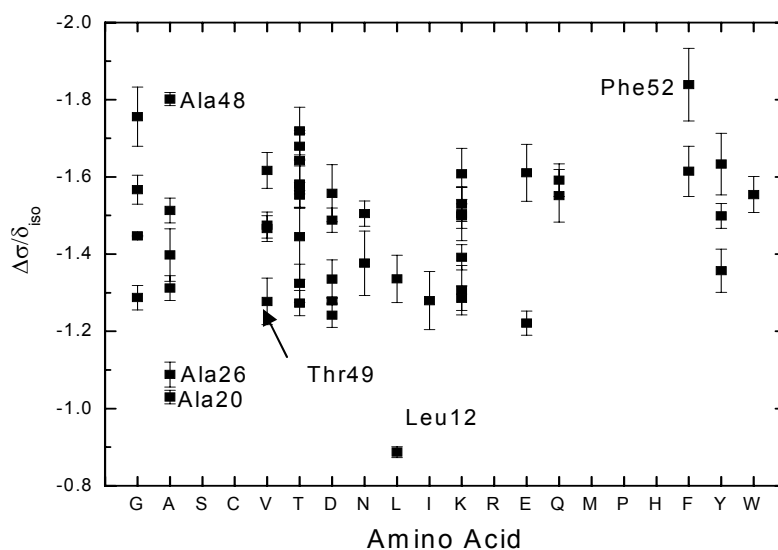
Once the site-specific values of the anisotropy have been determined, it is potentially interesting to examine the correlations (or lack thereof) between the measured chemical shielding anisotropies and structural parameters in the protein obtained from independent measurements. We observed no significant correlation between CSA values and the isotropic chemical shift (Figure 7.4.4). The Pearson's  $r^{117}$  describing the correlation is  $r=0.1$ , however, there is a slight tendency for nuclei

with lower-than-average  $^{15}\text{N}$  isotropic shifts to have lower-than-average  $^{15}\text{N}$   $|\Delta\sigma|$ s and for nuclei with higher-than-average isotropic  $^{15}\text{N}$  chemical shifts to have higher-than-average absolute values of the CSAs.

**Figure 7.4.4 CSAs in GB3 as a function of their isotropic chemical shifts (chemical shift scale shown from right to left as is conventional).** Some residues in GB3 with particularly high or low values of the isotropic shift are labeled in this plot.



**Figure 7.4.5 CSAs in GB3 as a function of amino acid type.**

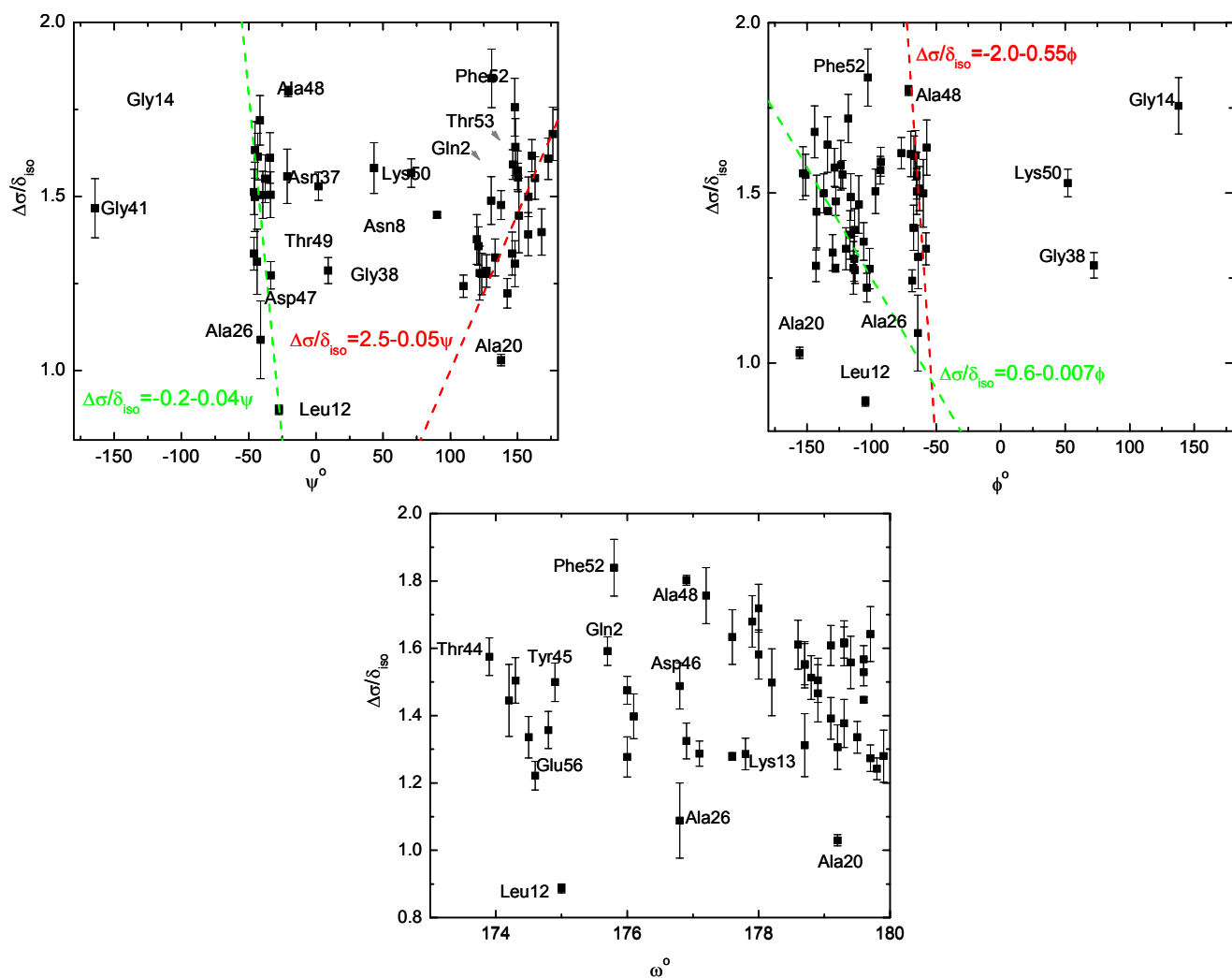


With amino acid type as well, there is no significant tendency (see Figure 7.4.5). Here  $\Delta\sigma/\delta_{\text{iso}}$  represents the chemical shielding anisotropy normalized by the isotropic chemical shift, and is plotted versus amino acid type. The correlations between the isotropic shifts and amino acid type, and nearest neighbor effects on isotropic shifts are well characterized<sup>157 158</sup>, so it was of interest to separate the dependence of the anisotropy of the shielding tensor on various parameters from the dependence of the isotropic chemical shift. Of note, the aromatic residues in GB3 (Phenylalanine, Tyrosine, and Tryptophan) all have higher-than-the-average absolute values of  $\Delta\sigma/\delta_{\text{iso}}$ .

The mean CSAs of residues in the  $\alpha$ -helix and  $\beta$ -strands are shown in Table 7.4.1. There is a weak correlation between the  $\beta_z$  angles and secondary structure, with slightly smaller angles in the  $\beta$ -strands (mean angle 18.9°) and turns (mean angle 19.1°) than in the helix (where the mean angle is 21.0°). Both the CSAs and  $\beta_z$  angles show smaller variation in the  $\alpha$ -helix (where the standard deviations in the CSA and the angle are 18.1 ppm and 3.1°, respectively) compared to the  $\beta$ -strands (18.6 ppm and 4.7°), and even larger variations were observed in the loops/turns (26.3 ppm and 7.5°), possibly consistent with significantly different electronic arrangement in the secondary structures.

Figure 7.4.6 depicts the correlation between  $\Delta\sigma/\delta_{\text{iso}}$  with the three dihedral angles:  $\phi$ ,  $\psi$ , and  $\omega$  in degrees. The correlation for the residues in the beta sheets is better than in the  $\alpha$ -helix, possibly reflecting that the ranges of both  $\phi$  and  $\psi$  values, as well as the  $\Delta\sigma$ s, in the beta sheets are considerably broader than the corresponding ranges in the  $\alpha$ -helix. There is no discernable correlation between  $\Delta\sigma/\delta_{\text{iso}}$  and  $\omega$ .

**Figure 7.4.6** Residue-specific values of the anisotropy of the chemical shielding tensor,  $\Delta\sigma$ , normalized by the isotropic chemical shift and plotted as a function of the backbone dihedral angles,  $\phi$ ,  $\psi$ , and  $\omega$  in degrees. Various residues are labeled with residue number and amino acid type for reference. The green and red dashed lines in the  $\phi$  and  $\psi$  plots correspond to the results of linear least-squares fits of the residues in the beta sheets (red) and  $\alpha$ -helix (green).



#### 7.4.5 Variability in the CSAs in GB3 Separated from Uncertainty

The observed range of site-specific  $^{15}\text{N}$  CSA values reflects both true CSA variability and random statistical errors in the measured parameters<sup>53</sup>. To address the actual variability of the CSA tensor we adopted the same statistical approach as in<sup>53, 54</sup> that assumes that the CSA values in proteins follow a Gaussian distribution. Assuming that the experimentally determined uncertainties are correct, the “true” values of the mean CSA ( $\mu$ , in ppm) and site-to-site CSA variability ( $\Lambda$ , also in ppm) can be determined by maximizing the following likelihood function<sup>54, 117</sup>:

$$p(\mu, \Lambda) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi(\Lambda^2 + (\delta\Delta\sigma_i)^2)}} \exp\left(-\frac{(\mu - \Delta\sigma_i)^2}{2(\Lambda^2 + (\delta\Delta\sigma_i)^2)}\right). \quad (7.4.1)$$

Here  $N$  is the number of residues probed in the measured distribution,  $\Delta\sigma_i$  and  $\delta\Delta\sigma_i$  are the measured CSA value and its experimental uncertainty for residue  $i$ .

The likelihood functions  $p(\mu, \Lambda)$  generated from the results of each of the three CSA determination methods are shown in Fig. 7.4.7. From the  $2R_2$ - $R_1$  method, the normally distributed CSA values in GB3 are characterized by a mean of  $\mu = -173.9$  ppm and the site-to-site variability  $\Lambda = 21.4$  ppm (see Eq. 7.4.1). We estimate a joint 95% confidence interval for  $\mu$  from this method to range from  $-165.7$  to  $-182.2$  ppm and for  $\Lambda$  from  $16.6$  to  $28.6$  ppm (Fig. 7.4.7). It is worth pointing out that a qualitative appreciation of the site-to-site variability in the CSA in GB3 is evident from a comparison of the linear dependence of  $2R_2' - R_1'$  on  $\omega_N^2$  (Eqs. 7.2.10-7.2.12) for three residues with similar  $J(0)$  values (Fig. 7.2.1). From the  $R/\eta$  method, the maximization of the likelihood function yielded the true variability in  $\Delta\sigma$  of  $\Lambda = 17.6$  ppm and a true mean CSA of  $-177.2$  ppm. We estimate a 95% confidence interval on

$\mu$  from this method to be from  $-169.9$  to  $-184.6$  ppm and for  $\Lambda$  from  $13.2$  to  $24.3$  ppm (Fig. 7.4.7).

The true mean CSA values ( $\mu$ ) from these methods are slightly higher in absolute value than those observed earlier in ubiquitin (mean CSA =  $-157$  ppm)<sup>51,52,57</sup> and in Rnase H ( $\mu = -172$  ppm)<sup>53</sup>, and slightly lower than those recently reported for ubiquitin<sup>54</sup> ( $\mu = -179.6$  ppm when scaled to a NH bond length of  $1.02$  Å), although within the average uncertainty of both of these measurements. The true site-to-site variability  $\Lambda$  in  $^{15}\text{N}$  CSA obtained here is comparable to the standard deviation of the CSA values in ubiquitin<sup>51,52</sup> but significantly bigger than the  $\Lambda$  values reported for Rnase H<sup>53</sup> and recently for ubiquitin<sup>54</sup>. The CSA distribution in ubiquitin, reconstructed from the individual CST components reported in<sup>56</sup>, is in a better agreement with our data for GB3: the standard deviations in these CSAs range from  $10.1$  to  $13.7$  ppm, and the site-to-site variability,  $\Lambda$ , from  $7.8$  to  $10.5$  ppm, depending on the model of local motion.

The value of  $\Lambda$  extracted from the observed site-specific CSA values, naturally, depends on the experimental uncertainties in CSA. Therefore, at least in principle, higher  $\Lambda$  values in GB3 could be a result of an underestimation of the experimental errors in the CSA. However, several lines of evidence suggest that this is not the case here. First of all, the residuals of fit from the diffusion tensor analyses (Table 7.1.1, rightmost column) are smaller than the ideal value of  $\chi^2/\text{df} \sim 1$ . This suggests that the errors in the relaxation and cross-correlation rates were possibly overestimated rather than underestimated. Second, the residuals of fit in the LS analysis (uniform CSA of  $-160$  ppm) of the autorelaxation data and NOEs at each

field separately passed the goodness-of-fit test for the overwhelming majority of residues in GB3 (98%, 96%, 100%, 98%, and 84% of residues passed the 95% confidence test at 9.4, 11.7, 14.1, 16.4 and 18.8 T, respectively, and 97% overall), also suggesting that the errors in the relaxation data were not underestimated. Third, in order to reduce  $\Lambda$  to the 5 ppm level reported in <sup>53,54</sup>, we had to scale up significantly the experimental errors in CSA (by a factor of 3 for the R/ $\eta$  method, 4 for the LS-CSA method, and >6.5 for the 2R<sub>2</sub>-R<sub>1</sub> method) assuming that all errors are uniformly underestimated. This scaling factor is too big, given the reasonable  $\chi^2/\text{df}$  values in all other fits presented here.

In addition, to further explore this issue, we introduced a certain  $\chi^2/\text{df}$  cutoff level (determined here by a 95% confidence level for the goodness-of-fit test <sup>117</sup>) as a highly conservative criterion for eliminating fits from consideration here. This cutoff excludes those residues where the robust regressions were acceptable but the  $\chi^2/\text{df}$  of the least-squares fit was too high due to an outlier that was effectively ignored by the robust methods: there are 9 such exclusions from the 2R<sub>2</sub>-R<sub>1</sub> method, 6 from R/ $\eta$  and 4 from the LS-CSA fit. The reasons for these outliers are unclear and do not appear to be systematic. For example, for the 800 MHz rates it seems possible that the different method of water suppression used in the R<sub>1</sub> and R<sub>2</sub> experiments may have affected the values of these rates for residues at the extreme edges of the spectra (e.g. Thr49 and Glu56), however this is clearly not the cause of the outliers at 400-700 MHz. Visual inspection of the spectral regions of the other outliers does not show any obvious or systematic deviations in baseline levels or lineshapes. If these outliers are removed, so that only those residues with the  $\chi^2/\text{df}$  of the least-squares fit lower than its 95%



confidence limit are considered (32 amides from the  $2R_2-R_1$  method, 33 from  $R/\eta$ , and 25 from LS-CSA, represented by the filled symbols in Figs. 7.4.1b,c,d, and Fig. 7.4.2b), the CSA variability from each method is reduced to what could probably be considered its lower bound in GB3:  $\Lambda_{2R_2-R_1}=10.6$  ppm,  $\Lambda_{R/\eta}=10.2$  ppm, and  $\Lambda_{LS-CSA}=11.9$  ppm. These estimates of the site-to-site CSA variability are still, consistently, almost a factor of two higher than those reported for Rnase H<sup>53</sup> or recently for ubiquitin<sup>54</sup>).

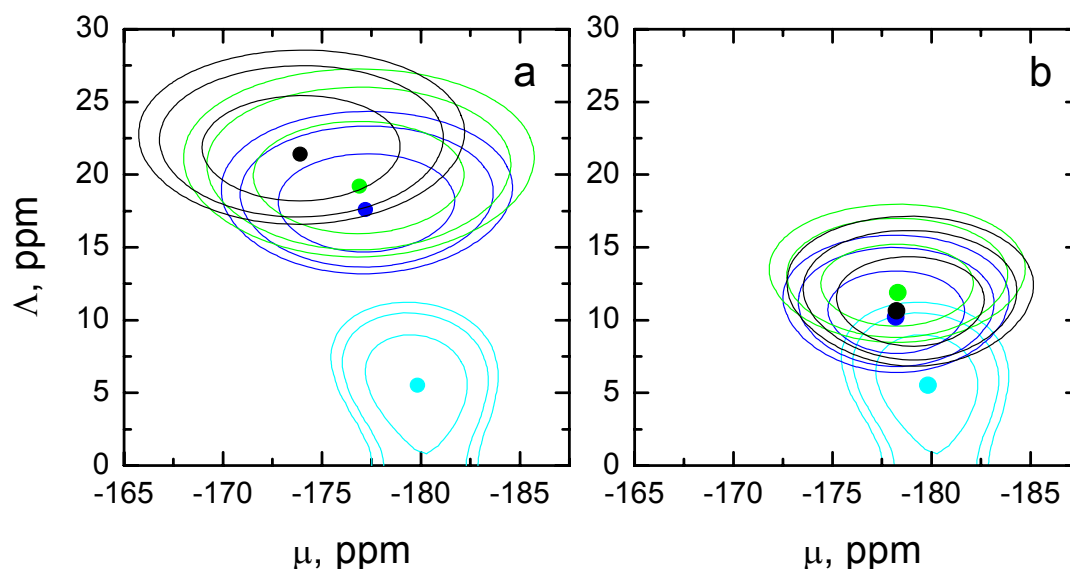
The results obtained here also differ from the  $^{15}\text{N}$  CSA statistics in short peptides, where for a set of 39 solid-state NMR data (summarized in<sup>34</sup>) we estimate a mean CSA of  $-155.8$  ppm and a standard deviation of the distribution of  $5.8$  ppm. The bigger range of CSA variability in GB3 compared to peptides could reflect greater internal structural heterogeneity in proteins.

To explore the effect of outliers as a possible source of the higher CSA variability observed here, we excluded from the set of residues for which  $p(\mu, \Lambda)$  was generated for each method the extrema of the corresponding CSA range (Fig.7.4.1a, Fig.7.4.3b). The mean CSA values were largely unchanged ( $\mu=-174.0$ ,  $-177.4$ , and  $-176.3$  ppm, for  $2R_2-R_1$ ,  $R/\eta$ , and LS-CSA, with Leu12 and Phe52, Ala26 and Phe52, and Gly38 and Phe52, excluded respectively) and the measures of the site-to-site variability  $\Lambda$  were reduced to  $17.2$ ,  $14.1$ , and  $13.3$  ppm, respectively. Restricting the CSA distribution even further by excluding all seven outliers in Fig.5.4.3 (Leu12, Ala20, Ala26, Gly38, Ala48, Thr49, and Phe52), thus effectively reducing the distribution to that contained within the dashed Gaussian curve shown in Fig. 7.4.3, reduced the calculated site-to-site CSA variability  $\Lambda$  to  $11.5$ ,  $13.8$ , and  $13.1$  ppm (for

2R<sub>2</sub>-R<sub>1</sub>, R/η, and LS-CSA, respectively), while the values of the true mean  $\mu$  were only slightly affected (−176.5, −177.9, and −176.3 ppm, respectively). These exclusions also resulted in similar changes for the distribution function generated from the average CSAs of the three methods (Fig. 7.4.3):  $\mu = -175.1$  ppm and  $\Lambda = 13.5$  ppm. Note that all these reduced estimates of the site-to-site variability in <sup>15</sup>N CSA are still significantly larger than those reported in<sup>53,54</sup>.

In summary, all these data then suggest that the site-to-site variability in <sup>15</sup>N CSA reported here for GB3 is most probably correctly estimated, or underestimated. This conclusion has important implications for the analysis of protein dynamics, since this degree of variability in the <sup>15</sup>N CSA means that the assumption of a uniform <sup>15</sup>N CSA value could result in significant errors in LS parameters.

**Figure 7.4.7 The likelihood functions (Eq. 7.4.1) obtained from different methods illustrate the significant site-to-site variability in the <sup>15</sup>N CSA values.** Contour plots of the likelihood functions  $p(\mu, \Lambda)$  (Eq. 7.4.1) corresponding to the <sup>15</sup>N CSA values from the three methods (2R<sub>2</sub>-R<sub>1</sub> (black), R/η (blue), and LS-CSA (green)) (a) for all analyzed residues in GB3 and (b) for only those residues where  $\chi^2/\text{df}$  from the least-squares fits passed the goodness-of-fit test at a 95% confidence level. Also shown (in cyan), for comparison, is the analogous likelihood function obtained for the recently reported <sup>15</sup>N CSAs in ubiquitin (Damberg, J. Am. Chem. Soc. 2005), scaled to a NH-bond length of 1.02 Å. The location of the maximum for each function is indicated by a dot (see also Table 7.4.1), the contour lines represent 68.3%, 90% and 95% bivariate confidence regions for  $\mu$  and  $\Lambda$ . In panel a, the 95% joint confidence intervals (in ppm) for  $\mu$  and  $\Lambda$  are (−165.7, −182.2) and (16.6, 28.6) from 2R<sub>2</sub>-R<sub>1</sub>, (−169.9, −184.6) and (13.2, 24.3) from R/η, and (−168.0, −185.7) and (14.3, 27.3) from LS-CSA methods. For a subset of residues (panel b) that pass the  $\chi^2/\text{df}$  cutoff, the corresponding confidence intervals for  $\mu$  and  $\Lambda$  are (−172.7, −185.2) and (6.8, 17.1) from 2R<sub>2</sub>-R<sub>1</sub>, (−172.5, −183.9) and (6.4, 15.8) from R/η, and (−171.8, −184.7) and (8.5, 18.0) from LS-CSA methods.



It is potentially interesting to examine the CSA variability obtained here in relationship to the spread in the isotropic chemical shifts in GB3. The isotropic chemical shift ( $\delta_{iso}$ ) and the CSA are both combinations of the principal values of the  $^{15}\text{N}$  CST:  $\delta_{iso} = (\delta_{xx} + \delta_{yy} + \delta_{zz})/3 \approx \sigma_{ref} - (\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3$ ;  $\Delta\sigma \approx \sigma_{zz} - (\sigma_{xx} + \sigma_{yy})/2$ , where  $\sigma_{ref}$  is the isotropic shielding of the reference compound, and the equation for  $\Delta\sigma$  used here is an approximate form of Eq.7.2.2, which is exact in the case of the axial symmetry of the CST, Eq.7.2.4. Assuming a random model, when all three components of the  $^{15}\text{N}$  CST are allowed to vary from site to site and are normally distributed with equal variances<sup>33</sup>, one can obtain from these equations the following relationship between the standard deviations in the CSA (here referred to as the variability  $\Lambda$ ) and in the isotropic chemical shift ( $\Delta\delta_{iso}$ ):

$$\Lambda = \frac{3}{\sqrt{2}} \kappa \cdot \Delta\delta_{iso}. \quad (7.4.2)$$

where  $\kappa$  is a numeric coefficient reflecting the interrelationship between the individual components of the CST:

$$\kappa = \sqrt{\frac{3 - 2R_{zx} - 2R_{zy} + R_{xy}}{3 + 2R_{zx} + 2R_{zy} + 2R_{xy}}} \quad (7.4.3)$$

Here  $R_{ij}$  is the correlation coefficient between  $\sigma_{ii}$  and  $\sigma_{jj}$ . In a particular case when all three CST components vary completely independently,  $\kappa = 1$ . Given the standard deviation of the isotropic chemical shift in GB3 is 6.5 ppm, the expected value of  $\Lambda$  in this case is 13.8 ppm. This number is smaller than the CSA variability obtained for all residues in GB3 ( $\Lambda_{2R2-R1}=21.4$ ,  $\Lambda_{R/\eta}=17.6$ , and  $\Lambda_{LS-CSA}=19.2$  ppm) but slightly larger than the values ( $\Lambda_{2R2-R1}=10.6$ ,  $\Lambda_{R/\eta}=10.2$ , and  $\Lambda_{LS-CSA}=11.9$  ppm) obtained

when considering only those residues with  $\chi^2/\text{df}$  below the 95% goodness-of-fit cutoff. The deviation in the value of  $\kappa$  from 1 suggests that the individual components of the  $^{15}\text{N}$  CST tensor are not independent from each other, however, it is impossible at this stage to draw a more definitive conclusion about the correlation coefficients between the individual components, and further studies are required to address this issue.

For example, it follows from Eq.7.4.2 that a positive correlation between  $\sigma_{xx}$  and  $\sigma_{yy}$ , both being independent of  $\sigma_{zz}$  will give  $\kappa < 1$  (with the lower bound at  $\kappa = 2/\sqrt{5}$ ), while an anti-correlation of these two components will result in  $\kappa > 1$  (up to  $\sqrt{2}$ ) with the upper bound on the CSA variability at  $\Lambda = 3 \Delta\delta_{\text{iso}}$  (or 19.5 ppm for GB3). It has been suggested<sup>159</sup> that  $\sigma_{xx}$  and  $\sigma_{yy}$  possibly vary in an anti-correlated manner – this would be consistent with the CSA variability in GB3 larger than 13.8 ppm. However, if the  $^{15}\text{N}$  CST is truly axially symmetric (i.e.  $\sigma_{xx} = \sigma_{yy}$ , hence  $R_{xy} = 1$ ), then the  $\Lambda$  value is expected to be smaller,  $\Lambda = 3\sqrt{2/5}\Delta\delta_{\text{iso}}$ , which gives the CSA variability around 12.3 ppm for GB3, again assuming that  $\sigma_{xx}$  and  $\sigma_{zz}$  (or  $\sigma_{\perp}$  and  $\sigma_{\parallel}$  in this case) are normally distributed and vary independently. A positive correlation between  $\sigma_{\perp}$  and  $\sigma_{\parallel}$  will further reduce the  $\Lambda$  values, down to zero at full correlation, while the anti-correlation will result in greater  $\Lambda$ s, with an upper bound at  $\Lambda = 6 \Delta\delta_{\text{iso}} = 39$  ppm. Using the correlation coefficients calculated from a collection<sup>34</sup> of 39 solid-state NMR data on short peptides,  $R_{zx} = 0.06$ ,  $R_{zy} = 0.43$ ,  $R_{xy} = -0.12$ , one would expect  $\Lambda$  of 14 ppm in GB3. Inserting into Eq.7.4.2 the correlation coefficients between the individual components of the  $^{15}\text{N}$  CST recently measured in ubiquitin<sup>56</sup>,

we estimate  $\Lambda$  to range from 9.6 to 13.3 ppm in ubiquitin (where the standard deviation in the isotropic chemical shift is 5.9 ppm) and from 10.5 to 14.6 in GB3.

### **7.5 Backbone Order Parameters in GB3 from Multifield $^{15}\text{N}$ Relaxation Data; The Effect of Site-Specific versus Uniform $^{15}\text{N}$ CSAs**

#### **7.5.1 Backbone Order Parameters: Assuming a Uniform $^{15}\text{N}$ CSA.**

When relaxation data ( $R_1$ ,  $R_2$ , NOE) at several fields are available, order parameters for a given NH vector can be obtained from the data at each field separately or from a simultaneous fit of the relaxation data for all available field strengths. Because the LS backbone dynamics should not depend on the applied magnetic field, all these order parameters are expected to agree with each other.

We first analyzed the relaxation data at each field separately using a standard LS approach<sup>152</sup> assuming a uniform value of  $^{15}\text{N}$  CSA of  $-160$  ppm. In all these analyses the quality of fit was very good: the residuals of the fit for the majority of residues (96% at 9.4T, 96% at 11.7 T, 98% at 14.1 T, 94% at 16.4 T, 84% at 18.8 T, and 94% overall) were within the acceptance level for a 95%-confidence goodness-of-fit test<sup>117</sup>, which indicates that the uncertainties in the experimental data are correct or overestimated. However, there is a striking discrepancy between the derived order parameters corresponding to different field strengths (Fig. 7.5.1a); for most residues in GB3 the observed variation in the derived  $S^2$  values among the fields exceeds their experimental uncertainties. Even in well-ordered parts of the protein, the difference in derived  $S^2$  between 800 and 400 MHz data exceeds 0.10 for some residues. Similar results were obtained when using  $^{15}\text{N}$  CSA of  $-170$  ppm or the mean

CSA of  $-174.2$  ppm (the mean CSA from the three determination methods,  $2R_2-R_1$ ,  $R/\eta$ , and LS-CSA). The observed disagreement between the derived  $S^2$  values obtained for the same NH group from the measurements at different fields thus raises significant concern about the accuracy of the order parameters derived by the standard analysis.

We also attempted to analyze simultaneously the relaxation data at all five fields using a uniform CSA of  $-160$  ppm and the average diffusion tensor and NH vector orientations from the crystal structure. This analysis indicated serious problems of fitting—only 8 out of 51 (Tyr3, Lys4, Leu5, Val6, Thr16, Ala23, Lys28 and Ala29) amides had residuals of the fit ( $\chi^2$ ) which passed the goodness-of-fit test at 95% confidence level <sup>117</sup>. Using a uniform CSA of  $-170$  ppm did not significantly improve the fit: here only 12 residues (Tyr3, Lys4, Leu5, Thr16, Thr18, Lys19, Ala23, Lys28, Gln32, Asp46, Thr51, and Thr55) had acceptable  $\chi^2$  values. Using the mean CSA value of  $-174.2$  ppm (see above) gave only 14 residues (Gln2, Tyr3, Leu5, Thr16, Thr18, Lys19, Ala23, Lys28, Gln32, Ala34, Val42, Tyr45, Asp46, and Thr51) with acceptable  $\chi^2$  values. These results from multiple approaches clearly indicate that the conventional LS approaches assuming a uniform  $^{15}\text{N}$  CSA fail to describe the multi-field experimental data in GB3.

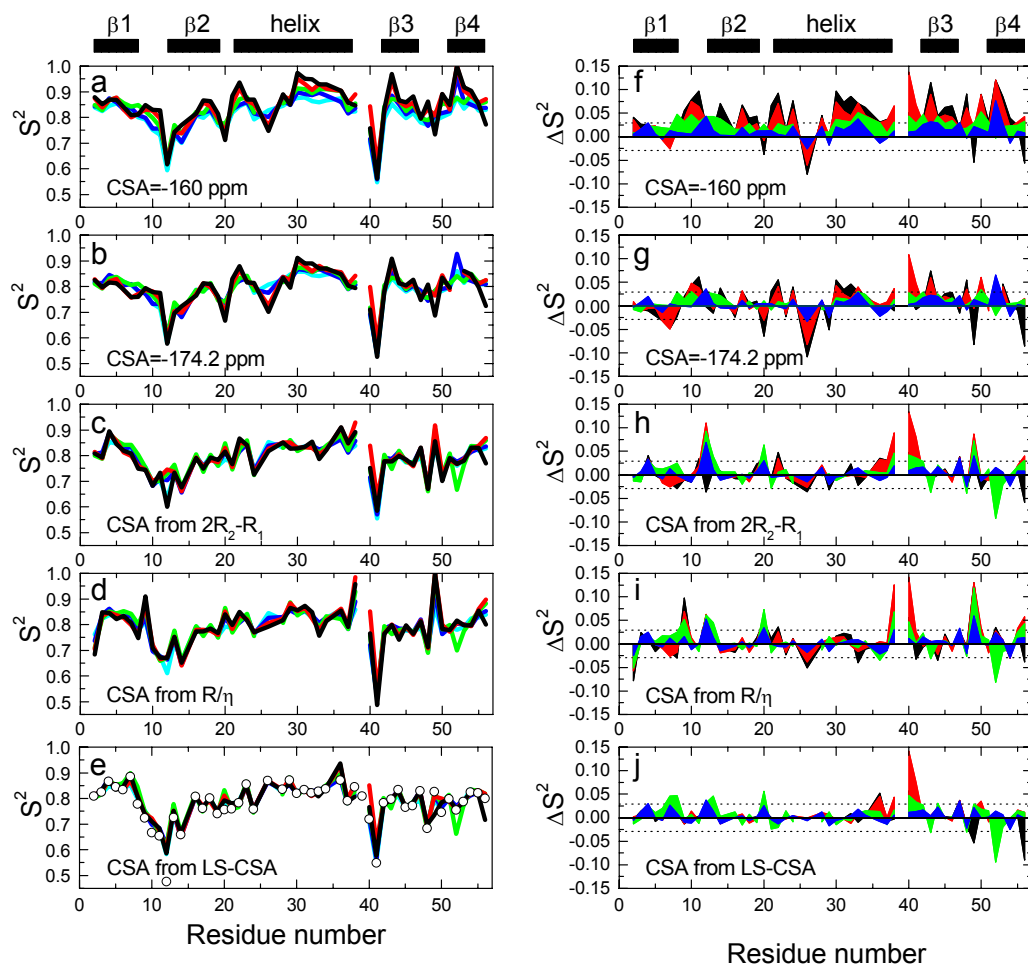
It is noteworthy that for most residues in GB3, the observed difference in the order parameters appears systematic, i.e. it increases with the field strength (Figs. 7.5.1a,e). This behavior could arise from conformational exchange contributions to  $^{15}\text{N}$   $R_2$  not accounted for in the analysis or deviations in the site-specific values of  $^{15}\text{N}$  CSA from their assumed values. Site-specific deviations in the  $^{15}\text{N}$ - $^1\text{H}$  bond length

from a uniform value of 1.02 or 1.04 Å could also result in erroneous order parameters; however, the currently available experimental data on variations in the NH bond length in proteins are insufficient to rigorously address this issue. A failure of the LS spectral density model to accurately represent data at multiple fields cannot be excluded (e.g. <sup>151,160</sup>), particularly with regard to the assumption of uncoupled local and global motions, however our analysis indicates that a modified LS model (using site-specific CSAs) fits the observed spectral densities in GB3.

Several lines of evidence suggest that conformational exchange is not the source of the observed discrepancy in the order parameters in GB3. As shown in Chapter 6, conformational exchange contributions are negligible for most of amides in GB3, except possibly Val39. This conclusion is also confirmed by the agreement (Fig. 7.1.1) between the measured  $R_2$ s and their reconstructed “exchange-free” values <sup>101</sup>,  $R_{2\text{free}}' = R_1' \cdot \eta_{xy} / \eta_z$ . The exclusion of conformational exchange as a possible cause of the observed discrepancy between the  $S^2$  values is further supported by the results of a LS analysis of the data at the individual fields. Here, 12 (excluding Val39) residues (Tyr3, Leu5, Ile7, Thr16, Ala23, Tyr33, Asp36, Asn37, Asp40, Thr44, Ala48, and Thr51) required a  $R_{\text{ex}}$ -containing model of local motion <sup>143</sup> at 18.8 T, where the  $R_{\text{ex}}$  contribution is expected to be the strongest. These  $R_{\text{ex}}$  values were relatively small (maximum  $0.53 \pm 0.10 \text{ s}^{-1}$  for Asp36 at 18.8 T) and likely reflect errors in LS model-selection, because the only residue that systematically showed conformational exchange at all five fields was Val39. In addition, excluding  $R_2$ s from the simultaneous analysis of the five-field data (hence using only  $R_1$ s and NOEs, as suggested in <sup>151</sup>) did not improve the quality of fit for CSA=-160 ppm: only 9

residues passed the goodness-of-fit test (Tyr3, Leu5, Lys13, Thr16, Lys19, Ala23, Ala29, Thr51, and Thr5) in this case. Note also that in terms of spectral densities, the presence of  $R_{\text{ex}}$  contribution will affect  $J(0)$  but not the  $J(\omega_N)$  values (Eqs 7.3.2, 7.3.3), hence the introduction of the  $R_{\text{ex}}$  terms might force the  $J(0)$  values from different fields to converge, but will not improve the fit of spectral densities at  $\omega=\omega_N$  (Fig.7.5.2) derived assuming a uniform CSA of  $-160$  ppm (see below). Finally, the  $R_{\text{ex}}$ -free values of overall diffusion tensor obtained solely from the cross-correlation measurements are in excellent agreement with those from the  $R_2/R_1$  ratio (Table 7.1.1).





**Figure 7.5.1** Shown are backbone order parameters in GB3 derived from a LS analysis of the  $^{15}\text{N}$  relaxation data ( $R_1$ ,  $R_2$ , NOE) at different fields (left panels). Right panels represent the differences,  $\Delta S^2 = S^2 - S^2(9.4\text{T})$ , between the  $S^2$  values at a particular field and at 9.4 Tesla, where the  $^{15}\text{N}$  CSA contribution to  $^{15}\text{N}$  relaxation rates is the weakest. (a, f) The LS analysis was performed in a conventional way, i.e. assuming a uniform CSA of  $-160$  ppm for all residues. (b, g) The LS analysis was performed assuming a uniform CSA of  $-174.2$  ppm (the average of the site-specific CSAs in GB3, see Table 1) for all residues. (c, h) Site-specific  $^{15}\text{N}$  CSA values from the  $2R_2-R_1$  method were used as input parameters. (d, i) Site-specific  $^{15}\text{N}$  CSA values from the  $R/\eta$  method were used as input parameters. (e, j) The LS analysis was performed for each field separately using the site-specific CSAs derived from the global fit (LS-CSA) of all five fields. Also shown as open circles in panel (d) are the order parameters from the global fit. The coloring is as follows: the 18.8T data are shown in black, 16.4 T in red, 14.1 T in green, 11.7 T in blue, and 9.4 T data in cyan. The dashed horizontal lines represent the average estimated level ( $\pm 0.029$ ) of the experimental uncertainty in  $\Delta S^2$ . Val39 has been removed from all panels because of the conformational exchange contribution (Hall, J. Biomol. NMR 2003). In order to exclude deviations in  $S^2$  due to a change in the model selection for different fields in a few residues, all data presented here were obtained assuming a model of local motion (model 2 in (Mandel, J. Mol. Biol., 1995), model “B” in (Fushman, J. Mol. Biol., 1997) that includes  $S^2$  and  $\tau_{\text{loc}}$  as fitting parameters. Our model-selection analysis showed that for the majority of residues in the secondary-structure elements of GB3 this was the preferred model (Hall, J. Biomol. NMR 2003). Allowing freedom in the model selection led to even greater discrepancies between the order parameters from different fields, which, however, exhibit the same behavior as shown here. As a measure of the discrepancy in order parameters, the rmsd from the average (over all five fields)  $S^2$  value for each method is 0.024 (panel a), 0.015 (b), 0.010 (c), 0.012 (d), and 0.009 (e), calculated for the secondary structure elements only.

### 7.5.2 Backbone Order Parameters: the Effect of Site-Specific $^{15}\text{N}$ CSAs.

To verify that the observed field-dependence in the order parameters (Fig.7.5.1a) could reflect site-specific variations in the  $^{15}\text{N}$  CSA unaccounted for in the conventional LS analysis, we performed the same derivation as above, this time using as input the site-specific  $^{15}\text{N}$  CSA values measured using the model-independent approaches. As shown in Figs. 7.5.1b,c,f,g, the inclusion of site-specific  $^{15}\text{N}$  CSA has dramatically reduced the variation in the order parameters among the fields, which is now within the level of experimental noise for most residues.

We therefore modified the LS analysis by including CSA as an additional fitting parameter (LS-CSA method, Materials and Methods). This resulted in a significant improvement in the quality of fit of the five-field data analysis for the majority of residues in GB3. For example, when the  $^{15}\text{N}$  CSA was allowed to vary in the LS-CSA method, the mean  $\chi^2/\text{df}$  for residues in the secondary structure dropped from 7.12 (for a uniform CSA of  $-160$  ppm) to a value of 0.92. All of the secondary structure residues except for Ala26 and Phe52 now have  $\chi^2/\text{df}$  low enough to pass the goodness-of-fit test at a 95% confidence level. Altogether, 47 out of 49 analyzed residues exhibited a decrease in  $\chi^2$  of the LS fit, and in 40 residues there is also a decrease in  $\chi^2/\text{df}$ . The residues where the  $\chi^2/\text{df}$  is not improved (Asn8, Leu12, Lys13, Thr16, Gly38, Asp40, Gly41, Asp47, and Thr49) are all in flexible regions of GB3 except for Thr16 for which the resulting CSA ( $-162.3$  ppm) is very close to  $-160$  ppm and the residuals of fit were already sufficiently low:  $\chi^2/\text{df}=0.56$  and  $0.67$  for the LS and LS-CSA methods, respectively.

For those residues where a reduction in  $\chi^2$  was accompanied by an increase in the number of fitting parameters (33 residues in GB3), a statistical F-test was performed<sup>117</sup> to determine if the improvement in the  $\chi^2$  was significant. For 31 (94%) of these residues, the reduction in the  $\chi^2$  is statistically justified at a 95% significance level or higher (i.e. the probability,  $P$ , that the reduction in  $\chi^2$  occurred by chance is  $P < 0.05$ ). For 25 (76%) of these residues the significance level is higher than 99% (i.e.  $P < 10^{-2}$ ), and for 22 (67%) of these residues the significance level is even higher than 99.9% (i.e.  $P < 10^{-3}$ ).

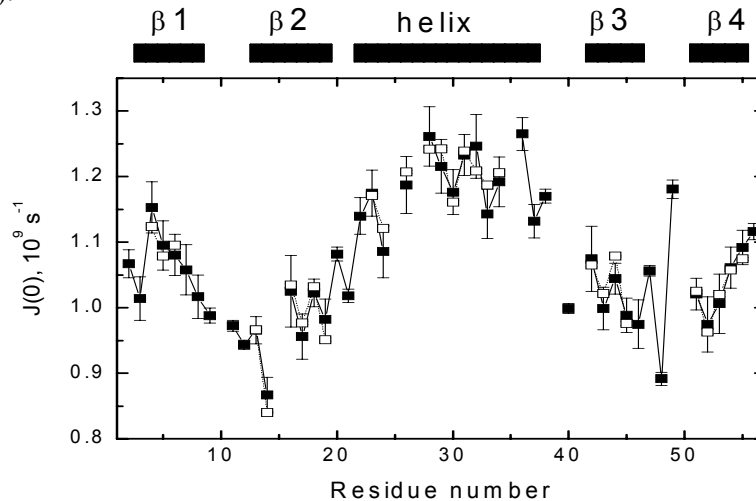
The order parameters derived from a simultaneous (global) fit of data from all five fields using the LS-CSA method are shown as open symbols in Fig 7.5.1d. All three regression methods had good agreement (within the experimental uncertainty for both  $S^2$  and the CSA) for 28 out of 49 amides in GB3 (Gln2 not included here because its coordinates are unavailable from the crystal structure). For an additional four residues (Gly9, Asp36, Asn37, and Gly41) the two robust methods agreed within their experimental uncertainties (68.3% confidence interval). For 17 residues (Tyr3, Ile7, Asn8, Lys10, Thr11, Leu12, Ala20, Val21, Asp22, Gly38, Asp40, Asp46, Asp47, Ala48, Thr49, Lys50, and the C-terminal Glu56), all of which are either in the loops/termini or at the edges of secondary structure elements, no CSA is reported here for the LS-CSA method because all three regression methods disagreed for either  $S^2$  or  $\Delta\sigma$ .

The “model-free” site-specific  $^{15}\text{N}$  CSA values were in the range from  $-126.0 \pm 3.9$  ppm (Ala26) to  $-243.4 \pm 4.7$  ppm (Phe52), with a mean of  $-176.9$  ppm, a median of  $-176.8$  ppm, and standard deviation of 20.0 ppm. The average estimated

level of the experimental errors is 1.76% (or 3.1 ppm) for the CSA, which gives a true site-to-site CSA variability  $\Lambda$  of 19.2 ppm and a true mean of  $-176.9$  ppm. We estimate a 95% confidence interval from this method to be from  $-167.7$  to  $-186.0$  ppm for  $\mu$  and from 14.2 to 27.6 ppm for  $\Lambda$  (Fig.7.4.7).

Using these site-specific  $^{15}\text{N}$  CSA values as input for the LS analyses at separate fields resulted in a further reduction in the spread of the order parameters among the fields (Figs. 7.5.1d,h). These results clearly indicate that the discrepancy in the order parameters in Fig.7.5.1a is caused by site-specific variations in the  $^{15}\text{N}$  CSA. To further validate the characteristics of the backbone dynamics ( $S^2$ ,  $\tau_{\text{loc}}$ ) derived simultaneously with site-specific  $^{15}\text{N}$  CSAs (LS-CSA method), we compared the spectral density  $J(\omega)$  at  $\omega=0$  reconstructed from these data with  $J(0)$  values obtained directly from the  $2\text{R}_2\text{-R}_1$  method (recall that this latter  $J(0)$  is *independent* of the  $^{15}\text{N}$  CSA). The good agreement between the two values of  $J(0)$  (Fig.7.5.2) for the secondary structure elements of GB3 thus validates the LS parameters derived using the LS-CSA method.

**Figure 7.5.2. The agreement between the spectral density component,  $J(0)$ , measured using the  $2\text{R}_2\text{-R}_1$  method and reconstructed from the LS parameters.** The spectral density component  $J(0)$  obtained from the  $2\text{R}_2\text{-R}_1$  method directly (solid symbols) and calculated from the order parameters and local correlation times obtained in the LS-CSA method (open symbols). Throughout this paper, the factor 2/5 arising from the normalization of the spectral density of the overall rotational diffusion is explicitly included in the corresponding expression for  $J(\omega)$ .

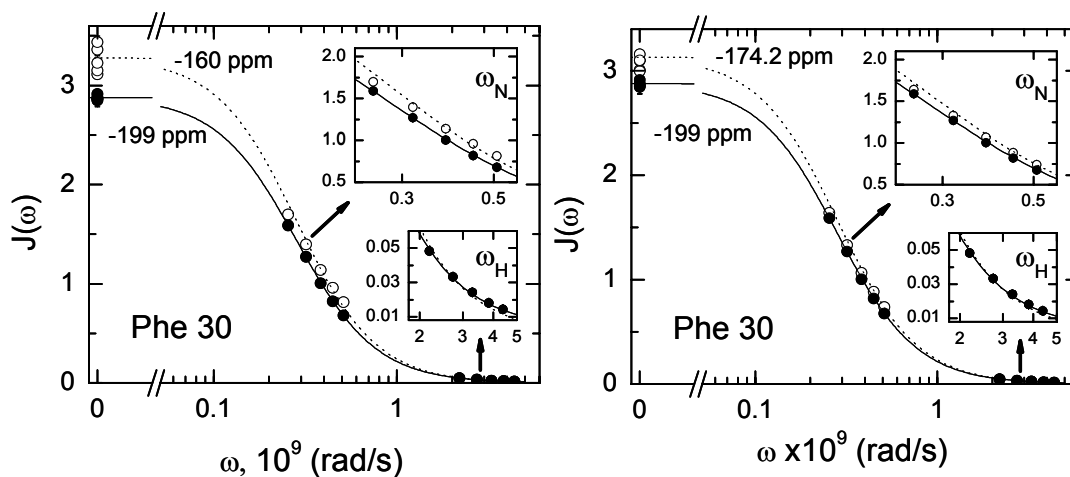


### 7.5.3 LS Fit of the Spectral Densities Directly.

A direct analysis of the spectral densities produced similar results. For a uniform CSA of  $-160$  ppm, the  $\chi^2/\text{df}$  of the fit of the spectral density functions at all five fields for the secondary structure elements of GB3 ranges from 0.46 (Tyr16) to 20.6 (Trp43) with a mean value of 4.73. The quality of the fits of the spectral density functions for Phe30 is shown in Fig.7.5.3. Overall, major discrepancies between the experimental data and the LS model were for  $\omega=0$ , due to the spread in the  $J(0)$  values derived at various fields, and at  $\omega=\omega_N$ , where the experimental  $J(\omega_N)$  values noticeably deviate from the theoretical curve. There is a good agreement for the high-frequency components (which are CSA-independent), particularly taking into account the reduced spectral density approximation<sup>79,78</sup> (Eq.7.3.1) made when deriving  $J(0.87\omega_H)$  from the experimental data.

The inclusion of CSA as a third fitting parameter (in addition to  $S^2$  and  $\tau_{\text{loc}}$ , see LS-SDF in Materials and Methods) resulted in the reduction of the residuals of fit for 29 out of 35 residues (or 83%) in the secondary structure elements; the  $\chi^2/\text{df}$  with this additional adjustable parameter ranged from 0.3 (Thr18) to 6.1 (Phe52) with a mean of 1.25. The LS-SDF method resulted in a significantly better convergence of  $J(0)$  values from different fields and, at the same time, in a better fit of the  $J(\omega_N)$  values (Fig.7.5.3). A similar improvement in the fit was obtained when using site-specific CSA values from the  $2R_2-R_1$  method, resulting in reduced  $\chi^2/\text{df}$  for 27 amides in the secondary structure.

**Figure 7.5.3.** Representative LS fit of all spectral density components from the five-field measurements for Phe30. Symbols depict the  $J(\omega)$  values for  $\omega = 0, \omega_N$ , and  $0.87\omega_H$  derived from relaxation data for each field separately (Eqs. 15-17) assuming CSA of  $-160$  ppm (open circles, left panel) or  $-174.2$  ppm (open circles, right panel) or the CSA value of  $-199.1$  ppm for Phe30 that optimizes the fit (solid circles both panels). The corresponding fit curves are shown as dashed and solid lines, respectively. Shown in the insets is a blow up of the regions corresponding to  $\omega = \omega_N$  and  $0.87\omega_H$ , indicated as “ $\omega_N$ ” and “ $\omega_H$ ”. The values of  $S^2$  and  $\tau_{loc}$  were  $0.93$  and  $3.0$  ps when using CSA of  $-160$  ppm,  $0.89$  and  $7.4$  ps when using CSA of  $-174.2$  ppm, and  $0.81$  and  $10.3$  ps for the fit CSA values. A 35-fold decrease in  $\chi^2/df$  was observed when using the CSA and the LS parameters from the LS-SDF fit compared to  $-160$  ppm, and a 12-fold decrease in  $\chi^2/df$  compared to when using  $-174.2$  ppm. The  $\Delta\sigma$  value derived using the  $2R_2-R_1$  method ( $-194.3$  ppm for Phe30) resulted in a fit which was practically indistinguishable from the LS-SDF fit shown here, as does the use of the CSA value ( $\Delta\sigma = -196.9 \pm 2.93$  ppm) from the LS-CSA fit for Phe30.



## 7.6 Discussion and Conclusions

### 7.6.1 Possible Sources of Systematic Errors in $^{15}\text{N}$ CSA Determination from Multiple-Field Data.

In addition to the imprecision in the CSA values caused by random noise associated with the measurements, there could be systematic errors – largely inaccuracy – stemming from the underlying assumptions in the analysis. Here we focus on some of them, a detailed analysis can be found elsewhere<sup>100</sup>.

*The N-H bond length.* As it is clear from Eqs. 7.2.9, 7.2.12, and 7.2.14, the  $^{15}\text{N}$  CSA values are determined via the dipolar term  $d$ , and therefore depend on our knowledge

of the NH-bond length. Here, we have, as is customary, assumed a uniform value of the NH bond. Site-to-site variations in  $r_{\text{HN}}$  will necessarily affect the  $\Delta\sigma$  values. However, the currently available information on the variations in the NH-bond length is insufficient for a rigorous analysis of this issue. Also, the CSA values derived here were obtained assuming the NH-bond length of 1.02 Å. For comparison with the CSA data obtained for  $r_{\text{HN}}=1.04$  Å, our results should be uniformly scaled by  $(1.02/1.04)^3=0.94$  (see also <sup>100</sup>). Thus, the mean <sup>15</sup>N CSA and the site-to-site variability (average of all three methods) obtained here correspond to –164.3 ppm and 20.0 ppm, respectively, if  $r_{\text{HN}}$  is 1.04 Å.

*The spectral densities.* The usual assumption made when analyzing <sup>15</sup>N relaxation data, be it LS approach or the model-independent analyses, is to neglect the difference between the spectral densities describing the effect of motion on the contributions to spin Hamiltonian from the <sup>15</sup>N-<sup>1</sup>H dipolar interaction ( $J_{\text{DD}}(\omega)$ ) and from the <sup>15</sup>N CSA ( $J_{\text{CSA}}(\omega)$ ), i.e.  $J_{\text{DD}}(\omega)=J_{\text{CSA}}(\omega)=J(\omega)$ . In general, however <sup>161,100</sup>,  $J_{\text{DD}}(\omega) \neq J_{\text{CSA}}(\omega)$ , and a correction for the difference between the spectral densities can be included as:

$$\Delta\sigma^{\text{correct}} = \Delta\sigma \cdot f, \quad (7.6.1)$$

where  $f$  is the correction factor:  $f = [J_{\text{DD}}(0)/J_{\text{CSA}}(0)]^{1/2}$  for the 2R<sub>2</sub>-R<sub>1</sub> method,  $f = \{[4J_{\text{DD}}(0) + 3J_{\text{DD}}(\omega_{\text{N}})]/[4J_{\text{CSA}}(0) + 3J_{\text{CSA}}(\omega_{\text{N}})]\}^{1/2}$  for R<sub>2</sub>/η<sub>xy</sub> and  $f = [J_{\text{DD}}(\omega_{\text{N}})/J_{\text{CSA}}(\omega_{\text{N}})]^{1/2}$  for R<sub>1</sub>/η<sub>z</sub>. There are several reasons why the spectral densities  $J_{\text{DD}}(\omega)$  and  $J_{\text{CSA}}(\omega)$  are not the same <sup>100</sup>. First, the nature of the chemical shielding suggests that it will fluctuate when the local environment of a nucleus changes as a result of internal motions in a protein. Not only the orientation (as usually assumed in

the equations relating relaxation rates to the spectral densities) of the CST but also the principal values themselves are expected to fluctuate. In contrast, the NH-bond length is less likely to change with motion, except possibly when transient hydrogen bonding occurs during protein dynamics. Note also that the changes in local environment that modulate the CST do not necessarily have to affect the orientation of the NH bond. A detailed analysis of the “breathing” of the  $^{15}\text{N}$  CST requires molecular dynamics simulations (e.g.<sup>162</sup>).

Second, even when neglecting the differences in the mechanisms of modulation of these two tensors by motions within a protein, the difference between the spectral densities is expected to arise from the fact that the CSA and dipolar tensors are not collinear. As follows from our data (Fig.7.4.2c), the average angle  $\beta_z$  between the NH vector and the z-axis of the CSA tensor is  $19.9^\circ$ . The effect of CSA-dipolar noncollinearity on the contribution to the spectral density from anisotropic overall tumbling has been analyzed in detail in<sup>161</sup>. Our calculations using the average site-specific CSAs from the three methods and the  $\beta_z$  angles (from  $R/\eta$ , Fig.7.5.2c) for GB3 resulted in the contributions from the noncollinearity to relaxation and cross-correlation rates that were on average within their respective experimental errors. As a result, the inclusion of these corrections in the model-independent and LS methods outlined above had no significant effect on the derived CSA values.

In addition, because of the anisotropic character of backbone motion in proteins<sup>163,164</sup>, where the principal mode of motion is rocking of the peptide plane about the  $\text{C}_\alpha\text{--C}_\alpha$  axis, the CSA-dipolar noncollinearity will result in different



amplitudes (and associated order parameters) of the NH-vector and CSA tensor motions. To investigate the effect of noncollinearity due to anisotropic backbone motions, we explored the difference in the order parameters for the NH vector and for a vector (representing the  $\sigma_{zz}$  axis) tilted by  $20^\circ$  towards the carbonyl atom in the peptide plane in a model system undergoing angular fluctuations about the  $C_\alpha-C_\alpha$  axis. We found that the maximum difference in the squared order parameters for these vectors was 5%, with  $S_{CSA}^2$  always smaller than  $S_{NH}^2$ , for a rotational angle of  $40^\circ$ , which is well above the maximum amplitude of Gaussian angular fluctuations about this axis recently reported for GB3<sup>148</sup>. Assuming that the correlation time of GAF motion is similar to that of the LS model, and that the order parameters are close to 1, Eq.7.6.1 gives  $f \approx S_{NH}/S_{CSA} < 1.03$ . This difference in the order parameters is insufficient to account for the large variability in the CSA that we observe in GB3. For example, if we assume for the sake of argument that the CSA in GB3 has a uniform value of  $-174.2$  ppm, the factor  $f$  would have to range from 0.7 to 1.6 (hence  $J_{DD}(0)/J_{CSA}(0)$  from 0.5 to 2.6) to account for the observed range of CSAs from the  $2R_2-R_1$  method. Similarly, to account for all the variability in the  $R/\eta$  measurements with respect to the average,  $f$  would have to vary from 0.7 to 1.4.

*The assumption of axial symmetry of the overall tumbling.* The order parameters and the  $^{15}\text{N}$  CSA values derived from the LS-based methods (but not those from the model-independent approaches) are sensitive to the model of overall tumbling used for the analysis. As presented here, the overall tumbling of GB3 in solution is anisotropic. While the axially symmetric and fully anisotropic tumbling models both provide a significant improvement in the fit over the isotropic diffusion model, the

axially symmetric model for the overall tumbling was assumed here, based on several lines of evidence.

(1) The molecular shape of GB3 to a good approximation is axially symmetric. The ratio of the principal values of the inertia tensor of the molecule is 1.80:1.79:1.00. Moreover, theoretical predictions for GB3 (Chapter 5) based on hydrodynamic calculations using HYDRONMR program gave a rotational diffusion tensor with ratio of principal components of 1.00:1.05:1.43, which suggests a high degree of axial symmetry.

(2) The fully anisotropic diffusion tensor derived from the relaxation data also shows a high degree of axial symmetry, with the principal values of the tensor,  $D_{xx}$  and  $D_{yy}$ , within their mutual errors at all fields. Also a global fit of the relaxation of data at all five fields resulted in a diffusion tensor with near zero rhombicity (0.08). This is also reflected in the large experimental uncertainties in the orientation of the x- and y-axes of the fully anisotropic tensor (angle  $\Psi$ ), indicating that the orientations of these axes of the diffusion tensor are not well defined.

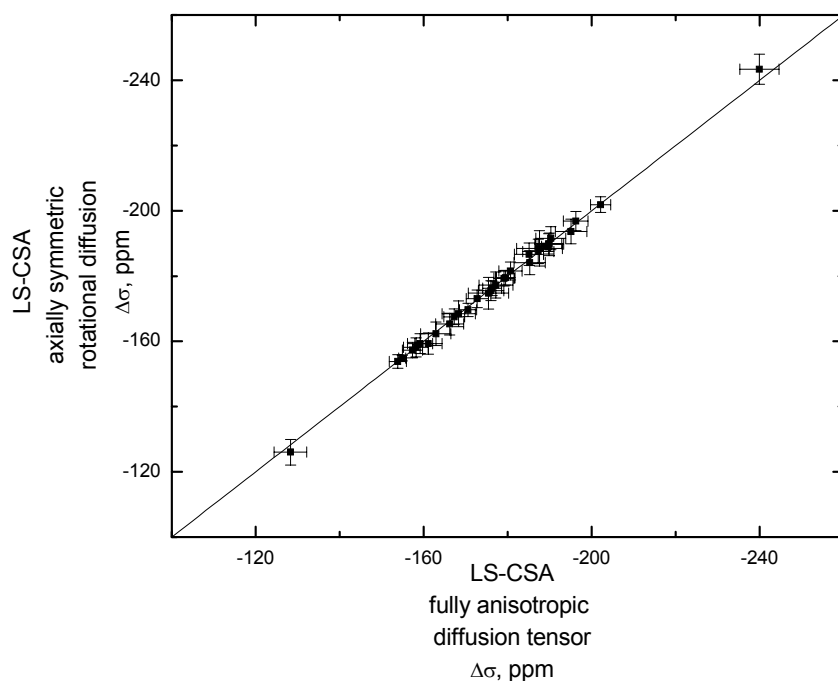
#### **<sup>15</sup>N CSAs and the order parameters: what errors in the order parameters to expect?**

As shown here, relaxation data at five spectrometer fields allowed an accurate assessment of the site-specific <sup>15</sup>N CSAs, and these values, in turn influenced the order parameters extracted from the data (Fig.7.5.1). Because measurements at multiple fields (particularly higher fields) are not always available to a general NMR user, it is instructional to estimate here the level of uncertainties in the order parameters expected from the use of a constant CSA instead of the true CSA values. A comparison of the order parameters obtained from the LS-CSA analysis of all five-

field data with those obtained for a typical field of 14.1 T, assuming a constant CSA, gave pair-wise rmsd values of 0.06 (or 6.5%, range of deviations from  $-0.06$  to  $0.11$ ) for  $-160$  ppm and 0.04 (or 4.1%, range from  $-0.09$  to  $0.07$ ) for  $-174.2$  ppm. The corresponding numbers for 11.4 T were, naturally, smaller: rmsd = 0.04 (4.9%, range from  $-0.04$  to  $0.09$ ) for  $-160$  ppm and 0.03 (3.2%, range  $-0.07$  to  $0.06$ ) for  $-174.2$  ppm. This comparison included only residues in the secondary structure of GB3, the deviations in the loop regions could be larger. Thus, even at low fields, the errors in the order parameters might not be negligible, particularly for those applications where quantitative changes in order parameters are of importance (as e.g. entropy changes monitored by  $^{15}\text{N}$  relaxation).

| Magnetic field (Tesla)                            | $D_{xx}^a$<br>( $10^7 \text{ s}^{-1}$ ) | $D_{yy}^a$<br>( $10^7 \text{ s}^{-1}$ ) | $D_{zz}^a$<br>( $10^7 \text{ s}^{-1}$ ) | $\Phi^{o\ b}$ | $\Theta^{o\ b}$ | $\Psi^{o\ b}$ | $\tau_c^c$<br>(ns) | Anisotropy <sup>d</sup> | rhombicity <sup>e</sup> | $\chi^2/\text{df}^f$ |
|---|---|---|---|---------------|-----------------|---------------|--------------------|-------------------------|-------------------------|----------------------|
| From auto- and cross-relaxation rate measurements |   |   |   |               |                 |               |                    |                         |                         |                      |
| 9.4   | 4.17(0.34)                              | 4.56(0.26)                              | 6.32(0.58)                              | 87(16)        | 60(12)          | 155(42)       | 3.33(0.16)         | 1.45(0.15)              | 0.30(0.06)              | 0.61                 |
| 11.7  | 4.28(0.16)                              | 4.68(0.28)                              | 6.22(0.49)                              | 94(12)        | 77(16)          | 141(28)       | 3.29(0.13)         | 1.39(0.12)              | 0.34(0.06)              | 0.58                 |
| 14.1  | 4.39(0.18)                              | 4.54(0.32)                              | 6.03(0.67)                              | 91(13)        | 76(19)          | 125(52)       | 3.34(0.17)         | 1.35(0.16)              | 0.14(0.03)              | 0.74                 |
| 16.4  | 4.35(0.19)                              | 4.55(0.24)                              | 6.17(0.40)                              | 101(12)       | 62(9)           | 95(51)        | 3.32(0.11)         | 1.39(0.10)              | 0.18(0.03)              | 0.88                 |
| 18.8  | 4.44(0.82)                              | 4.49(0.67)                              | 6.15(2.21)                              | 101(20)       | 66(10)          | 117(259)      | 3.31(0.54)         | 1.38(0.52)              | 0.04(0.03)              | 0.79                 |
|   | 4.41                                    | 4.50                                    | 6.11                                    | 97            | 67              | 104           | 3.33               | 1.37                    | 0.08                    | 0.71                 |
| From cross-correlation rate measurements          |   |   |   |               |                 |               |                    |                         |                         |                      |
| 9.4   | 4.28(0.33)                              | 4.66(0.24)                              | 5.99(0.40)                              | 97(20)        | 67(18)          | 130(63)       | 3.35(0.13)         | 1.34(0.11)              | 0.37(0.08)              | 0.61                 |
| 11.7  | 4.24(0.41)                              | 4.48(0.32)                              | 6.22(0.85)                              | 90(14)        | 55(11)          | 141(77)       | 3.34(0.22)         | 1.43(0.21)              | 0.19(0.05)              | 0.94                 |
| 14.1  | 4.35(0.25)                              | 4.44(0.20)                              | 6.24(0.66)                              | 93(9)         | 65(6)           | 169(53)       | 3.33(0.16)         | 1.42(0.16)              | 0.07(0.02)              | 0.54                 |

**Figure 7.6.1. CSAs determined using the LS-CSA method assuming the fully anisotropic rotational diffusion tensor (given in Table 7.6.1) plotted versus CSAs determined using the LS-CSA method assuming the axially symmetric rotational diffusion tensor (given in Table 7.1.1).**



## 7.6.2 Conclusions

This is a comprehensive study of the  $^{15}\text{N}$  chemical shielding anisotropy in the GB3 domain based on a combination of  $^{15}\text{N}$  relaxation and CSA/dipolar cross-correlation measurements at five static magnetic fields. The analysis was performed using various combinations of the experimental data and using model-independent approaches as well as methods based on Lipari-Szabo approximation. The results indicate significant site-to-site variations in the principal values and the orientation of the  $^{15}\text{N}$  CSA, similar to those observed earlier in ubiquitin<sup>51,52</sup>. Our estimates of the

true variability in the  $^{15}\text{N}$  CSA in GB3 depend to some degree upon which method for determining the CSA was used and which subset of residues is considered. These estimates range from 10.2 ppm (for the 33 residues that pass the  $\chi^2/\text{df}$  cutoff from the  $R/\eta$  method) to 21.4 ppm for all 47 residues from the  $2R_2-R_1$  method. Although this range of values could be a result of limited statistics, all of these estimates are still larger than the derived variability in the  $^{15}\text{N}$  CSA from studies of ribonuclease H<sup>53</sup> or recently of ubiquitin using a subset of the methods used here<sup>54</sup>. The true mean CSA values range from  $-173.9$  ppm ( $2R_2-R_1$ ) to  $-177.2$  ppm ( $R/\eta$ ). Our data show that using the site-specific values of the  $^{15}\text{N}$  chemical shielding anisotropy obtained here significantly improves the agreement between LS order parameters measured at different fields and allows simultaneous fit of the  $^{15}\text{N}$  relaxation data at five fields to LS spectral densities. These findings emphasize the necessity of taking into account the variability of the  $^{15}\text{N}$  chemical shielding tensor for accurate analysis of protein dynamics from  $^{15}\text{N}$  relaxation measurements. This can be achieved by including CSA as an additional fitting parameter in the LS analysis of multiple-field data, provided the sample temperature and other experimental conditions are the same at all fields/spectrometers. These analyses also show that the Lipari-Szabo form of the spectral density provides a satisfactory approximation for the experimental spectral densities obtained using a reduced spectral density approach.

## Chapter 8: Novel Solution-State Relaxation Measurements

### 8.1 Motivation

During the last two decades, during which much progress has been made towards eliminating the size limitation on molecules which can be studied by NMR, much work has focussed on pulse-sequences observing the proton ( $^1\text{H}$  nucleus), the most sensitive among the nuclei present in biological molecules. Due to its large gyromagnetic ratio ( $\gamma_{\text{H}}=2.675 \times 10^8 \text{ (T}\cdot\text{s)}^{-1}$ ), a proton undergoes the strongest interactions and has the fastest transverse relaxation—the former is an advantage in terms of sensitivity (protons are easy to excite) while the latter is a disadvantage (proton signals decay rapidly, and are broadened, especially in large molecules). Recent advances in NMR probe technology have made so-called “direct-detection” experiments on heteronuclei (e.g.  $^{13}\text{C}$  and  $^{15}\text{N}$ ) increasingly practical. Here we demonstrate the utility of direct-detected  $^{15}\text{N}$  relaxation experiments to measure relaxation in  $^{15}\text{N}\text{-}\{^2\text{H}\}$  groups in proteins.

As discussed previously here, and in <sup>30,78</sup>, derivation of protein dynamics from NMR  $^{15}\text{N}$  spin relaxation rates is an underdetermined problem: there are a multitude of interaction and motional parameters (dipolar coupling constants, chemical shift tensors, characteristics of the overall and local motions) that have to be obtained from a few experiments. The techniques of spectral density function mapping (developed by <sup>30</sup>) and reduced spectral density function mapping <sup>78 79</sup> allow characterization of local motions of N-H bond vectors from  $^{15}\text{N}$  relaxation data without assumptions about the specific form of the spectral density function—however, the spectral density function is only sampled by  $^{15}\text{N}\text{-}\{^1\text{H}\}$  relaxation rates at a limited number of

specific frequencies ( $J(\omega_N)$   $J(\omega_H)$  and  $J(\omega_N \pm \omega_H)$ ) so derivation of all of the parameters (for an extended Lipari-Szabo model there is the possibility of 7:  $S^2$ ,  $\tau_{loc}$ ,  $S^2_{fast}$ ,  $\tau_{fast}$ ,  $R_{ex}$ ,  $\Delta\sigma$ , and  $r_{NH}$ ) which describe both the motion ( $S^2$ - $R_{ex}$ ) and the structure of the residue-specific spin system ( $\Delta\sigma$  and  $r_{NH}$ ) is problematic.

As discussed in Chapter 7, there is strong evidence that the  $^{15}\text{N}$  CSA varies from site to site in GB3, therefore the site-specific CSA values are essential for an accurate picture of protein dynamics. While  $^{15}\text{N}$  CSA values can be obtained by combining relaxation measurements at several fields (as was described in Chapter 7), the accuracy and precision of such analysis could be limited by the low CSA contribution to  $^{15}\text{N}$  relaxation at the magnetic fields currently available. For example, the CSA contributes only 12% of the  $^{15}\text{N}$  transverse relaxation rate at 9.4 T, and even at 21.1 T its contribution is only about 41% (for a protein with a rotational correlation time of 5 ns). It has also been postulated<sup>165</sup> that the N-H bond distance,  $r_{NH}$ , might vary from site-to-site within a protein, possibly with a strong dependence on hydrogen bonding. Thus far, there has been no successful method for the separation of these effects (variation of the  $\Delta\sigma$  and variation of  $r_{NH}$ ) in relaxation rates or in their effects on derived dynamic parameters.

Here we use heteronuclear-detected experiments to measure  $^{15}\text{N}$  relaxation rates in proteins in  $\text{D}_2\text{O}$ , where the dipolar contribution is lessened by replacing the dipolar-coupled partner ( $^1\text{H}$ ) with a deuteron. Although the fast NH bond librations might be affected by the  $^1\text{H} \rightarrow ^2\text{H}$  replacement, the motions of the peptide planes and larger-scale segmental motions—the major contributors to NMR-detected local dynamics in the backbone<sup>32,11</sup>—are expected to be much less sensitive to the



deuteration. Because  $^{15}\text{N}$  CSA is the major mechanism of  $^{15}\text{N}$  relaxation in the ND spin pair, the proposed measurements could provide a more sensitive probe for determining  $^{15}\text{N}$  CSA tensors. Because these rates have less relative sensitivity to the dipolar coupling, they are less sensitive to variation in the NH bond length.

In addition, this method has the potential of extending the current repertoire of spectral densities sampled by  $^{15}\text{N}$  relaxation measurements, as the relaxation rates in the ND spin system are sensitive to motional averaging at different frequencies than in the NH pair. Specifically, relaxation rates in  $^{15}\text{N}\text{-}\{^2\text{H}\}$  systems sample the spectral density function at  $J(\omega_{\text{N}})$ ,  $J(\omega_{\text{D}})$ , and  $J(\omega_{\text{N}}\pm\omega_{\text{D}})$ , and it has been shown<sup>25</sup> that at these frequencies the spectral density function describing motion of the  $^{15}\text{N}\text{-}^2\text{H}$  bond is in agreement with what is predicted from the sampling of that of the  $^{15}\text{N}\text{-}^1\text{H}$  bond<sup>25</sup>. This means that a combination of relaxation rates in  $^{15}\text{N}\text{-}\{^1\text{H}\}$  and  $^{15}\text{N}\text{-}\{^2\text{H}\}$  systems samples the spectral density function at 8 frequency-points ( $J(0)$ ,  $J(\omega_{\text{N}})$ ,  $J(\omega_{\text{D}})$ ,  $J(\omega_{\text{H}})$ , and  $J(\omega_{\text{N}}\pm\omega_{\text{D}})$ ,  $J(\omega_{\text{N}}\pm\omega_{\text{D}})$ ) with two points ( $J(\omega_{\text{N}})$  and  $J(0)$ ) sampled by rates from both systems. Of these values,  $J(\omega_{\text{N}}+\omega_{\text{D}})$  is particularly interesting, because due to the opposite signs of  $\omega_{\text{N}}$  and  $\omega_{\text{D}}$ , it is at a particularly low frequency (26, 31, and 36 MHz at proton resonance frequencies of 500, 600 and 700 MHz) and is expected to be sensitive to motions in the ns range<sup>25</sup>.

The lower sensitivity of nitrogen detection compared to hydrogen detection can be compensated, at least in part, by the use of specially designed probes<sup>166</sup>. In addition, the  $^{15}\text{N}$  signals are significantly sharper than  $^1\text{H}$  signals, which partially recovers losses in sensitivity. When fast-relaxing systems are studied, it becomes increasingly common to make use of the early techniques employing heteronuclei for

the starting magnetization and the signal detection <sup>167</sup>. Recently, it has been shown that pulse sequences using <sup>13</sup>C detection are more sensitive than the classical <sup>1</sup>H-detected ones in the presence of fast relaxation induced by a paramagnetic center <sup>168</sup>. This has triggered the development of <sup>13</sup>C detection methods for applications to large proteins <sup>169,170,171</sup>.

**Table 8.1.1.** Percent contribution to <sup>15</sup>N R<sub>1</sub> and R<sub>2</sub> in <sup>15</sup>N-<sup>1</sup>H and <sup>15</sup>N-<sup>2</sup>H spin systems at 800 MHz from dipolar (terms which contain J<sub>DD</sub>(ω)) and CSA (J<sub>CSA</sub>(ω)) relaxation for a protein with a tumbling time of 5 ns. It has been assumed that the rates (R<sub>1</sub> and R<sub>2</sub>) can be written as sums of these terms, i.e. R<sub>1</sub> = Γ<sub>DD</sub><sup>z</sup> + Γ<sub>CSA</sub><sup>z</sup> and R<sub>2</sub> = Γ<sub>DD</sub><sup>xy</sup> + Γ<sub>CSA</sub><sup>xy</sup>. No quadrupolar relaxation terms are included here, since deuterium decoupling was applied throughout the relaxation delay.

| τ <sub>c</sub> =5ns, 800 MHz (18.8 T) | Γ <sub>DD</sub> <sup>z</sup><br>(% of R <sub>1</sub> ) | Γ <sub>CSA</sub> <sup>z</sup><br>(% of R <sub>1</sub> ) | Γ <sub>DD</sub> <sup>xy</sup><br>(% of R <sub>2</sub> ) | Γ <sub>CSA</sub> <sup>xy</sup><br>(% of R <sub>2</sub> ) |
|---------------------------------------|--|---|---|--|
| <sup>15</sup> N- <sup>1</sup> H}      | 64   | 36  | 64  | 36   |
| <sup>15</sup> N- <sup>2</sup> H}      | 42   | 58  | 15  | 85   |

**Table 8.1.2.** Calculated contributions to  $^{15}\text{N}$   $R_1$  and  $R_2$  in  $^{15}\text{N}$ - $\{^1\text{H}\}$  and  $^{15}\text{N}$ - $\{^2\text{H}\}$  spin systems in GB3 at five fields from various relaxation mechanisms. There are no quadrupolar relaxation terms included here, since deuterium decoupling was applied throughout the relaxation delay.

|  | 400 MHz            | 500 MHz            | 600 MHz            | 700 MHz            | 800 MHz            |
|--|--------------------|--------------------|--------------------|--------------------|--------------------|
|  | $\text{s}^{-1}$    | $\text{s}^{-1}$    | $\text{s}^{-1}$    | $\text{s}^{-1}$    | $\text{s}^{-1}$    |
| $R_1$ $^{15}\text{N}$ - $\{^2\text{H}^{\text{N}}\}^{\text{a}}$   | 0.85               | 0.72               | 0.61               | 0.52               | 0.45               |
| $R_1$ $^{15}\text{N}$ - $\{^1\text{H}^{\text{M}}\}^{\text{b}}$   | 0.09 (0.01)        | 0.06 (0.01)        | 0.05 (0.01)        | 0.04 (0.01)        | 0.03 (0.00)        |
| $R_1$ $^{15}\text{N}$ - $\{^2\text{H}^{\text{S}}\}^{\text{c}}$   | 0.87               | 0.74               | 0.63               | 0.54               | 0.46               |
| $R_1$ $^{15}\text{N}$ - $\{^{13}\text{C}^{\text{M}}\}^{\text{d}}$  | 0.15 (0.01)        | 0.11 (0.01)        | 0.08 (0.01)        | 0.06 (0.01)        | 0.05 (0.00)        |
| $R_1$ $^{15}\text{N}(\text{CSA})^{\text{e}}$   | 0.43               | 0.52               | 0.58               | 0.62               | 0.66               |
| $R_1$ (total) $^{\text{f}}$  | 1.54               | 1.43               | 1.34               | 1.27               | 1.20               |
| <b><math>R_1</math> (w/o <math>^{13}\text{C}</math>)<math>^{\text{g}}</math></b>                                     | <b>1.39</b>        | <b>1.33</b>        | <b>1.26</b>        | <b>1.20</b>        | <b>1.15</b>        |
| <b><math>R_1</math> (w/o <math>^{13}\text{C}</math>, <math>^2\text{H}^{\text{S}}</math>)<math>^{\text{e}}</math></b> | <b>1.36</b>        | <b>1.30</b>        | <b>1.24</b>        | <b>1.19</b>        | <b>1.14</b>        |
| <i>Measured</i>  |                    | <i>1.17 (0.11)</i> | <i>1.10 (0.16)</i> | <i>1.06 (0.23)</i> |                    |
| $R_2$ $\text{N}$ - $\{^2\text{H}^{\text{N}}\}^{\text{a}}$  | 0.82               | 0.72               | 0.65               | 0.59               | 0.54               |
| $R_2$ $\text{N}$ - $\{^1\text{H}^{\text{M}}\}^{\text{b}}$  | 0.16 (0.02)        | 0.15 (0.02)        | 0.14 (0.02)        | 0.14 (0.02)        | 0.13 (0.02)        |
| $R_2$ $\text{N}$ - $\{^2\text{H}^{\text{S}}\}^{\text{c}}$  | 0.87               | 0.78               | 0.70               | 0.64               | 0.59               |
| $R_2$ $\text{N}$ - $\{^{13}\text{C}^{\text{M}}\}^{\text{d}}$   | 0.19 (0.02)        | 0.16 (0.01)        | 0.14 (0.01)        | 0.13 (0.01)        | 0.12 (0.01)        |
| $R_2$ $\text{N}(\text{CSA})^{\text{e}}$  | 0.85               | 1.25               | 1.72               | 2.26               | 2.87               |
| $R_2$ (total) $^{\text{f}}$  | 2.07               | 2.34               | 2.71               | 3.17               | 3.72               |
| <b><math>R_2</math> (w/o <math>^{13}\text{C}</math>)<math>^{\text{g}}</math></b>                                     | <b>1.89</b>        | <b>2.18</b>        | <b>2.56</b>        | <b>3.04</b>        | <b>3.60</b>        |
| <b><math>R_2</math> (w/o <math>^{13}\text{C}</math>, <math>^2\text{H}^{\text{S}}</math>)<math>^{\text{h}}</math></b> | <b>1.83</b>        | <b>2.12</b>        | <b>2.51</b>        | <b>2.98</b>        | <b>3.54</b>        |
| <i>Measured</i>  |                    | <i>2.03 (0.21)</i> | <i>3.05 (0.44)</i> | <i>2.91 (0.42)</i> |                    |
|  |                    |                    |                    |                    |                    |
| $R_1$ $^{15}\text{N}$ - $\{^1\text{H}^{\text{N}}\}^{\text{a}}$   | 3.22               | 2.57               | 2.07               | 1.68               | 1.39               |
| $R_1$ $^{15}\text{N}$ - $\{^1\text{H}^{\text{M}}\}^{\text{b}}$   | 0.10 (0.01)        | 0.08 (0.01)        | 0.06 (0.01)        | 0.05 (0.01)        | 0.04 (0.01)        |
| $R_1$ $^{15}\text{N}$ - $\{^{13}\text{C}^{\text{M}}\}^{\text{d}}$  | 0.15 (0.01)        | 0.12 (0.01)        | 0.09 (0.01)        | 0.07 (0.01)        | 0.06 (0.01)        |
| $R_1$ $^{15}\text{N}(\text{CSA})^{\text{e}}$   | 0.43               | 0.54               | 0.63               | 0.70               | 0.76               |
| $R_1$ (total) $^{\text{f}}$  | 3.90               | 3.31               | 2.86               | 2.51               | 2.25               |
| <b><math>R_1</math> (w/o <math>^{13}\text{C}</math>)<math>^{\text{g}}</math></b>                                     | <b>3.74</b>        | <b>3.19</b>        | <b>2.77</b>        | <b>2.44</b>        | <b>2.19</b>        |
| <i>Measured</i>  | <i>2.96 (0.21)</i> | <i>2.57 (0.17)</i> | <i>2.26 (0.14)</i> | <i>2.04 (0.15)</i> | <i>1.87 (0.16)</i> |
| $R_2$ $\text{N}$ - $\{^1\text{H}^{\text{N}}\}^{\text{a}}$  | 5.15               | 4.80               | 4.54               | 4.33               | 4.18               |
| $R_2$ $\text{N}$ - $\{^1\text{H}^{\text{M}}\}^{\text{b}}$  | 0.16 (0.02)        | 0.15 (0.02)        | 0.14 (0.02)        | 0.13 (0.02)        | 0.13 (0.02)        |
| $R_2$ $\text{N}$ - $\{^{13}\text{C}^{\text{M}}\}^{\text{d}}$   | 0.17 (0.02)        | 0.15 (0.01)        | 0.13 (0.01)        | 0.12 (0.01)        | 0.11 (0.01)        |
| $R_2$ $\text{N}(\text{CSA})^{\text{e}}$  | 0.71               | 1.04               | 1.43               | 1.86               | 2.35               |
| $R_2$ $^{\text{f}}$  | 6.18               | 6.13               | 6.23               | 6.45               | 6.77               |
| <b><math>R_2</math> (w/o <math>^{13}\text{C}</math>)<math>^{\text{g}}</math></b>                                     | <b>6.01</b>        | <b>5.99</b>        | <b>6.10</b>        | <b>6.33</b>        | <b>6.66</b>        |
| <i>Measured</i>  | <i>4.78 (0.42)</i> | <i>4.81 (0.42)</i> | <i>5.02 (0.45)</i> | <i>5.28 (0.51)</i> | <i>5.58 (0.67)</i> |

Numbers in parentheses represent standard deviations over all residues and are given for structure-dependent parameters. Rates were calculated with the assumption of isotropic overall diffusion (though the anisotropy of GB3 is a significant contribution to the variation of  $^{15}\text{N}$  relaxation rates of individual residues, the averages-over-all residues of the rates are fairly insensitive to the anisotropy) and of no local motion ( $S^2=1$ ) with spectral density function given by:

$$J(\omega) = \frac{2}{5} \frac{\tau_c}{1 + (\omega\tau_c)^2}, \text{ where } \tau_c \text{ is the rotational correlation time.}$$

<sup>a</sup> Contribution to  $^{15}\text{N}$  relaxation from dipolar interaction with directly bound hydron, bond distance assumed to be 1.02 Å.

<sup>b</sup> Contribution to  $^{15}\text{N}$  relaxation from dipolar interactions with all other protons in the molecule (coordinates generated using X-plor on crystal structure 1IGD.pdb).

<sup>c</sup> Result given in a, corrected for impurity of the solution. In  $\text{D}_2\text{O}$  solution we estimate 3%  $^1\text{H}$  content. The relaxation for exchangeable groups (such as backbone amides) in a solvent which is a mixture of species is bi-exponential:  $(1 - f_i)e^{-R_s t} + f_i e^{-R_i t}$  where  $f_i$  is the volume fraction of the impurity,  $R_s$  is the relaxation rate when bound to the majority solvent constituent (given in a), and  $R_i$  is the relaxation rate when bound to the impurity. In a mostly  $\text{D}_2\text{O}$  solution which contains some  $^1\text{H}$  impurity, fitting the bi-exponential relaxation to a mono-exponential function results in an overestimation of  $R_s$ .

<sup>d</sup> Mean-over-residues of the contribution to  $^{15}\text{N}$  relaxation from dipolar interactions with all  $^{13}\text{C}$  nuclei in the protein assuming 100% isotopic enrichment.

<sup>e</sup> Contribution to  $^{15}\text{N}$  relaxation rate from  $^{15}\text{N}$  chemical shielding anisotropy (assumed to be -170 ppm).

<sup>f</sup> Total  $^{15}\text{N}$  relaxation rate (sum of contributions b-e for  $^{15}\text{N}$ - $^2\text{H}$  rates and a-b and d-e for  $^{15}\text{N}$ - $^1\text{H}$  rates).

<sup>g</sup> Total relaxation rate in  $^{15}\text{N}$ -only labeled proteins (sum of contributions b, c, and e for  $^{15}\text{N}$ - $^2\text{H}$  rates and a, b, and e for  $^{15}\text{N}$ - $^1\text{H}$  rates).

<sup>h</sup> For  $^{15}\text{N}$ - $\{^2\text{H}\}$  measurements in  $^{15}\text{N}$  only labeled proteins with non-Inept filter (sum of contributions a, b, and e).

## 8.2 What is the Effect of $^1\text{H} \rightarrow ^2\text{H}$ Isotopic Substitution on the CSA of the $^{15}\text{N}$ Nucleus in a Peptide? A Quantum Mechanical Study

### **8.2.1 Expectation Based on the Observed Effect on the $^{15}\text{N}$ Isotropic Chemical Shift**

A simple “back of the envelope” calculation of the expected magnitude of the change in the  $^{15}\text{N}$  shielding anisotropy upon isotopic substitution of the directly bonded proton with a deuteron,  $\left| \Delta \Delta \sigma^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H}) \right| = \left| \Delta \Delta \delta^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H}) \right|$ , based on the size of the shifts in the isotropic chemical shifts,  $\Delta \delta_{iso}^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H})$ , assuming axial symmetry of the  $^{15}\text{N}$  CSA tensors for both  $^1\text{H}$  and  $^2\text{H}$  bound nuclei is shown below:

$$\begin{aligned} \Delta \delta_{iso}^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H}) &= \delta_{iso}^{15\text{N}}(^2\text{H}) - \delta_{iso}^{15\text{N}}(^1\text{H}) \\ \Delta \delta_{iso}^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H}) &= \frac{1}{3}(\delta_{\parallel}^{15\text{N}}(^2\text{H}) + 2\delta_{\perp}^{15\text{N}}(^2\text{H})) - \frac{1}{3}(\delta_{\parallel}^{15\text{N}}(^1\text{H}) + 2\delta_{\perp}^{15\text{N}}(^1\text{H})) \\ &= \frac{1}{3}(\delta_{\parallel}^{15\text{N}}(^2\text{H}) - \delta_{\parallel}^{15\text{N}}(^1\text{H})) + \frac{2}{3}(\delta_{\perp}^{15\text{N}}(^2\text{H}) - \delta_{\perp}^{15\text{N}}(^1\text{H})) \\ &= \frac{1}{3}\Delta \delta_{\parallel}^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H}) + \frac{2}{3}\Delta \delta_{\perp}^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H}), \end{aligned}$$

$\Delta \delta_{iso}^{15\text{N}}(^1\text{H} \rightarrow ^2\text{H})$  is known to be small ( $< 1$  ppm in proteins<sup>172</sup>), since this quantity is two orders of magnitude smaller than  $\left| \Delta \delta^{15\text{N}}(^1\text{H}) \right| \cong 170$  ppm, it can be approximated as zero relative to this parameter, therefore the above relation gives:

$\Delta\delta_{\parallel}^{15N}(^1H, ^2H) \cong -2\Delta\delta_{\perp}^{15N}(^1H \rightarrow ^2H)$ , so that

$$\begin{aligned}\Delta\Delta\delta^{15N}(^1H \rightarrow ^2H) &= (\delta_{\parallel}^{15N}(^2H) - \delta_{\perp}^{15N}(^2H)) - (\delta_{\parallel}^{15N}(^1H) - \delta_{\perp}^{15N}(^1H)) \\ &= (\delta_{\parallel}^{15N}(^2H) - \delta_{\parallel}^{15N}(^1H)) - (\delta_{\perp}^{15N}(^2H) - \delta_{\perp}^{15N}(^1H)) \\ &= \Delta\delta_{\parallel}^{15N}(^1H \rightarrow ^2H) - \Delta\delta_{\perp}^{15N}(^1H \rightarrow ^2H), \text{ can be approximated as:}\end{aligned}$$

$$\Delta\Delta\delta^{15N}(^1H \rightarrow ^2H) \cong \frac{3}{2}\Delta\delta_{\parallel}^{15N}(^1H \rightarrow ^2H) \cong -3\Delta\delta_{\perp}^{15N}(^1H \rightarrow ^2H),$$

so that finally:  $|\Delta\delta_{\perp}^{15N}(^1H \rightarrow ^2H)| < |\Delta\delta_{\parallel}^{15N}(^1H \rightarrow ^2H)| < |\Delta\Delta\delta^{15N}(^1H \rightarrow ^2H)|$ .

In  $^1H^N$ -N systems, the parallel component of the CST is the least shielded (most high-field) component, i.e.  $|\delta_{\parallel}^{15N}(^1H)| > |\delta_{\perp}^{15N}(^1H)|$ . In the event of  $^1H \rightarrow ^2H$  isotope substitution, it has been shown in other NH systems<sup>173,174</sup> that the anharmonicity of the bond potential energy and the larger deuteron mass result in a shortening of the effective  $^2H^N$ -N distance as compared to the  $^1H^N$ -N distance. This results in a higher electron density in the N- $^2H$  bond compared to the N- $^1H$  bond, which corresponds to a small increase in the  $^{15}N$  isotropic shielding. There is no available experimental information on the effect of  $^1H \rightarrow ^2H$  isotope substitution on the individual components of the  $^{15}N$  CST or on the anisotropy in the literature, however, it is expected from the geometry of the peptide backbone that the absolute value of the effect on the  $\delta_{\perp}$  components (i.e.  $|\Delta\delta_{\perp}^{15N}(^1H \rightarrow ^2H)|$  and/or  $|\Delta\delta_{yy}^{15N}(^1H \rightarrow ^2H)|$  and

$|\Delta\delta_{xx}^{15N}(^1H \rightarrow ^2H)|$  will be very small. If these changes are indeed negligible, it follows from above that,  $\delta_{\parallel}^{15N}(^1H \rightarrow ^2H) = 3\Delta\delta_{iso}^{15N}(^1H \rightarrow ^2H)$ , so  $\Delta\Delta\delta^{15N}(^1H \rightarrow ^2H) = 3\Delta\delta_{iso}^{15N}(^1H, ^2H)$ . As mentioned above,  $\Delta\delta_{iso}^{15N}(^1H \rightarrow ^2H)$  in proteins is  $< 1$  ppm, so  $|\Delta\Delta\delta^{15N}(^1H \rightarrow ^2H)|$  (and necessarily  $|\Delta\delta_{iso}^{15N}(^1H \rightarrow ^2H)|$  which is equivalent) should be  $< 3$  ppm, which is well within the estimated experimental errors of the measurement of  $\Delta\sigma^{15N}(^1H)$  from solution relaxation measurements.

Within the limits of the Born-Oppenheimer approximation, where the motions of the nuclei are very slow compared to the motion of the electrons, the electronic properties (e.g. chemical shielding) can be calculated with the nuclear positions being treated as fixed parameters. It should be feasible to perform an *ab initio* calculation of the effect of isotope substitution on the individual components of the  $^{15}N$  CST, including the effect of changes in the N- $\{^{1,2}H\}$  bond length. Stretching of the  $^1H^N$ -N bond, caused by hydrogen bonding in a peptide has previously been studied by *ab initio* calculation. These effects on the CSA are small,  $\sim 3$ -9 ppm<sup>47,49</sup>, when compared with the dependence of the  $^{15}N$  CSA values on backbone conformation from a similar calculation, which was  $\sim 30$ -40 ppm<sup>47</sup>. A recent quantum mechanical calculation study of a series of model dipeptides and Ala-X and X-Ala sequences (where X is any amino acid) in both  $\alpha$ -helical and  $\beta$ -sheet conformations showed that the principal values of the tensor were significantly affected by hydrogen bonding at both the carbonyl group and the N-H bond (which have been shown to stretch the N-H bond), by the adjacent residues in the polypeptide sequence, and by backbone conformation

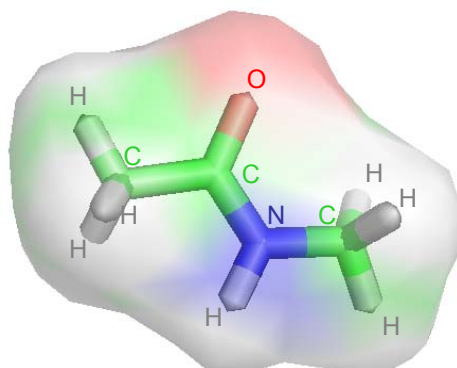
<sup>48</sup>. The magnitudes of the changes in the orientation of the tensor due to the effects were found to be insignificant compared to the changes in the principal values.

### 8.2.2 *Ab Initio* Quantum Mechanical Calculations

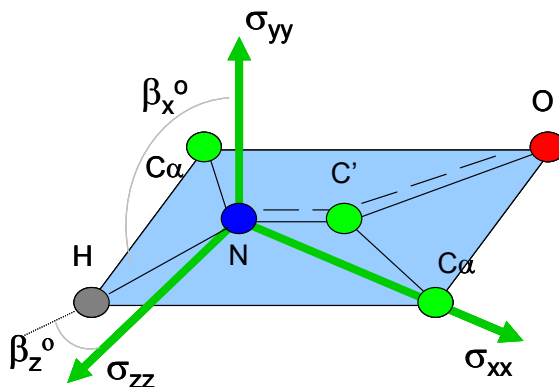
In order to confirm the prediction of our “back of the envelope” calculation that the change in the  $^{15}\text{N}$  chemical shift anisotropy due to isotopic substitution of the proton by a deuteron would be within the errors of our CSA measurement methods, we performed several quantum mechanical *ab initio* calculations of the optimized geometry and chemical shift tensors of a protonated and deuterated (at the N site) form of the N-methylacetamide molecule in the gas phase. The optimized geometry of N-methylacetamide using a restricted Hartree-Fock calculation is shown in Figure 8.2.1. We chose this molecule because it is the simplest model for the amide linkage of peptides and proteins. Previous *ab initio* studies of the  $^{15}\text{N}$  chemical shift tensor have shown both the principal values and orientations of this tensor to be sensitive to the way in which the solvent is modeled <sup>175</sup>. However, to a good approximation, the calculated tendencies in these parameters are independent of solvent model <sup>48</sup>. Since we are primarily interested in the *difference* in the anisotropy of the tensor between the states with bound proton and bound deuteron, and not in the absolute value of the components of the tensor in either case, we believe we can use the results of calculation in the gas phase.



**Figure 8.2.1.** Geometry of the N-methylacetamide (NMA) molecule at 0 Kelvin in the gas phase as calculated using a restricted Hartree-Fock optimization with a 6-311+G(2d,p) basis set. All calculations were repeated for both isotopomers of NMA (the molecule shown here and the molecule in which the amide proton was substituted with a deuterium).



**Figure 8.2.2.** Schematic illustration of the expected orientation of the principal components of the  $^{15}\text{N}$  CSA tensor (green) with respect to the peptide plane (light blue).  $\sigma_{zz}$  corresponds to the least shielded (most high field) component of the tensor, and lies approximately in the peptide plane, tilted from the NH bond by in-plane angle  $\beta_z$ .  $\sigma_{yy}$  is the next least shielded component and is expected to be roughly orthogonal to the peptide plane (i.e.  $\beta_x \sim 90^\circ$ ), while  $\sigma_{xx}$  is the most shielded component and lies in the peptide plane, perpendicular to  $\sigma_{zz}$ .



### *Restricted Hartree-Fock Calculation*

The first calculation we did was a restricted Hartree-Fock calculation using a 6-31G(2d,p) basis set. For  $\text{CH}_3\text{-}^{15}\text{N}^1\text{H-CO-CH}_3$  this calculation predicted a vibrationally averaged N-H bond length of 0.983 Å, a  $^{15}\text{N}$  isotropic chemical shift,  $\delta_{\text{iso}}$  of 193.0 ppm, and a  $^{15}\text{N}$  chemical shielding anisotropy,  $\Delta\sigma$  of 128.5 ppm.

Nitrogen chemical shifts are referenced to 0 ppm proton, based on the method of Wishart et al.<sup>176</sup>. As expected, none of the values of these parameters are similar to what we experimentally observe in proteins in solution at 297 K (where  $r_{\text{NH}}=1.02\text{-}1.04\text{ \AA}$ ,  $\delta_{\text{iso}}=100\text{-}135\text{ ppm}$ ,  $\Delta\sigma = 100\text{-}300\text{ ppm}$ ). For  $\text{CH}_3\text{-}^{15}\text{N}^2\text{H-CO-CH}_3$  the  $^{15}\text{N}\text{-}^2\text{H}$  bond length was  $0.984\text{ \AA}$ , the  $^{15}\text{N}$   $\delta_{\text{iso}}$  was  $192.9\text{ ppm}$ , and  $\Delta\sigma = 127.4\text{ ppm}$ . For this calculation, therefore, the  $^{15}\text{N}\text{-}^2\text{H}$  bond length was  $0.002\text{ \AA}$  longer than the  $^{15}\text{N}\text{-}^1\text{H}$  bond length and the  $^{15}\text{N}$  nucleus in  $^{15}\text{N}\text{-}\{^1\text{H}\}$  was  $0.1\text{ ppm}$  more shielded than in  $^{15}\text{N}\text{-}\{^2\text{H}\}$ , and the difference in anisotropy,  $\Delta\Delta\sigma^{15\text{N}}(^1\text{H}\rightarrow^2\text{H})$  was  $-1.1\text{ ppm}$  (the  $^{15}\text{N}$  shielding in  $^{15}\text{N}\text{-}\{^2\text{H}\}$  was slightly less anisotropic). Furthermore,  $\Delta\Delta\sigma^{15\text{N}}(^1\text{H}\rightarrow^2\text{H})$  from this calculation was primarily due to changes in the shielding of the  $\sigma_{zz}$  and  $\sigma_{xx}$  components of the tensor upon deuteration (in agreement with our intuitive predictions) with almost no change in  $\sigma_{yy}$ , the principal component which is orthogonal to the peptide plane. In  $\text{CH}_3\text{-}^{15}\text{N}^1\text{H-CO-CH}_3$  the principal values of the shielding tensor were:  $\sigma_{xx}=103.59$   $\sigma_{yy}=196.75$   $\sigma_{zz}=278.65$ ; in  $\text{CH}_3\text{-ND-CO-CH}_3$  they were  $\sigma_{xx}=104.18$   $\sigma_{yy}=196.74$   $\sigma_{zz}=277.85$ . Therefore  $\Delta\sigma_{xx}(^1\text{H}\rightarrow^2\text{H})=0.59\text{ ppm}$ ,  $\Delta\sigma_{yy}(^1\text{H}\rightarrow^2\text{H})=-0.01\text{ ppm}$ , and  $\Delta\sigma_{zz}(^1\text{H}\rightarrow^2\text{H})=-0.80\text{ ppm}$ . Though the relative magnitudes of the changes in the principal components agree with our expectation from the geometry of the molecule, the changes in the CSA and bond length are in the opposite directions (here  $\Delta\Delta\sigma^{15\text{N}}(^1\text{H}\rightarrow^2\text{H})$  is negative and the  $^{15}\text{N}\text{-}^2\text{H}$  bond length is longer than the  $^{15}\text{N}\text{-}^1\text{H}$  bond length) as our expectation (section 8.2.1).

The geometry optimization for this calculation took 3 hours, the frequency calculation took 1 day and 16 hours, and the calculation of the NMR chemical shifts

of  $^{15}\text{N}$  and  $^1\text{H}$  and  $^2\text{H}$  took 15 minutes. All calculations were performed on a Dual-CPU 2.5 GHz Apple G5 using the program *Gaussian* and initiated by graduate student Merle Zimmermann in the laboratory of Dr. John Tossell at UMD.

### *B3LYP Calculation*

The second calculation we did used a B3LYP6-311+G(2d,p) approach (this 6-311+G(2d,p) basis set is larger than that used in the Hartree-Fock calculation). For  $\text{CH}_3\text{-}^{15}\text{N}^1\text{H-CO-CH}_3$  this calculation predicted a vibrationally averaged N- $^1\text{H}$  bond length of 0.998 Å, a  $^{15}\text{N}$  isotropic chemical shift,  $\delta_{\text{iso}}$  of 153.97 ppm, and a  $^{15}\text{N}$  chemical shielding anisotropy,  $\Delta\sigma$  of 115.11 ppm. For  $\text{CH}_3\text{-}^{15}\text{N}^2\text{H-CO-CH}_3$  the N- $^2\text{H}$  bond length was 1.006 Å, the  $^{15}\text{N}$   $\delta_{\text{iso}}$  was 153.75 ppm, and  $\Delta\sigma = 114.23$  ppm. For this calculation, the vibrationally averaged N- $^2\text{H}$  bond length was 0.008 Å longer than the N- $^1\text{H}$  bond length and the  $^{15}\text{N}$  nucleus in  $^{15}\text{N-}\{^1\text{H}\}$  was 0.22 ppm more shielded than in  $^{15}\text{N-}\{^2\text{H}\}$ , with  $\Delta\Delta\sigma^{^{15}\text{N}}(^1\text{H}\rightarrow^2\text{H})=-0.88$  ppm. In  $\text{CH}_3\text{-}^{15}\text{N}^1\text{H-CO-CH}_3$  the principal values of the shielding tensor were:  $\sigma_{\text{xx}}=66.84$   $\sigma_{\text{yy}}=164.36$   $\sigma_{\text{zz}}=230.71$ ; in  $\text{CH}_3\text{-}^{15}\text{N}^2\text{H-CO-CH}_3$  they were  $\sigma_{\text{xx}}=67.33$   $\sigma_{\text{yy}}=164.01$   $\sigma_{\text{zz}}=229.90$ , therefore the changes in the components were all of roughly similar magnitude:  
 $\Delta\sigma_{\text{xx}}(^1\text{H}\rightarrow^2\text{H})=0.49$  ppm,  $\Delta\sigma_{\text{yy}}(^1\text{H}\rightarrow^2\text{H})=-0.35$  ppm, and  $\Delta\sigma_{\text{zz}}(^1\text{H}\rightarrow^2\text{H})=-0.81$  ppm. In general, this calculation predicts that all changes in all parameters (both the bond length and chemical shielding) to be so small as to be negligible to NMR relaxation experiments. The calculation time was similar to the time for the Hartree-Fock calculation.

### *MP2 Calculation*

The third and final calculation we did of the  $^{15}\text{N}$  chemical shifts in the NMA molecule was an MP2 calculation using the 6-311+G(2d,p) basis set. For  $\text{CH}_3\text{-}^{15}\text{N}^1\text{H-CO-CH}_3$  this calculation predicted a vibrationally averaged N- $^1\text{H}$  bond length of 0.969 Å, a  $^{15}\text{N}$  isotropic chemical shift,  $\delta_{\text{iso}}$  of 168.8 ppm, and a  $^{15}\text{N}$  chemical shielding anisotropy,  $\Delta\sigma$  of 146.8 ppm. For  $\text{CH}_3\text{-}^{15}\text{N}^2\text{H-CO-CH}_3$  the N- $^2\text{H}$  bond length was 0.960 Å, the  $^{15}\text{N}$   $\delta_{\text{iso}}$  was 166.2 ppm, and  $\Delta\sigma = 154.6$  ppm. For this calculation, therefore, the N- $^2\text{H}$  bond length was 0.009 Å *shorter* on average than the N- $^1\text{H}$  bond length and the  $^{15}\text{N}$  nucleus in  $^{15}\text{N}\text{-}\{^1\text{H}\}$  was 2.6 ppm more shielded than in  $^{15}\text{N}\text{-}\{^2\text{H}\}$ , with  $\Delta\Delta\sigma^{^{15}\text{N}}(^1\text{H}\rightarrow^2\text{H})=7.8$  ppm. These changes are in the opposite direction to the changes in these parameters from the previous two calculations, however the direction of the changes predicted here (shorter bond length and greater shielding for  $^{15}\text{N}\text{-}\{^2\text{H}\}$ ) are in agreement with experimental data on other NH containing compounds<sup>173,174</sup>. In  $\text{CH}_3\text{-}^{15}\text{N}^1\text{H-CO-CH}_3$  the principal values of the shielding tensor were:  $\sigma_{\text{xx}}=64.95$   $\sigma_{\text{yy}}=174.88$   $\sigma_{\text{zz}}=266.72$ ; in  $\text{CH}_3\text{-}^{15}\text{N}^2\text{H-CO-CH}_3$  they were  $\sigma_{\text{xx}}=51.08$   $\sigma_{\text{yy}}=178.15$   $\sigma_{\text{zz}}=269.24$ . Therefore, the largest change was in  $\Delta\sigma_{\text{xx}}(^1\text{H}\rightarrow^2\text{H})=-13.87$  ppm, followed by  $\Delta\sigma_{\text{yy}}(^1\text{H}\rightarrow^2\text{H})= 3.27$  ppm, and  $\Delta\sigma_{\text{zz}}(^1\text{H}\rightarrow^2\text{H})=2.53$  ppm. This calculation took significantly more time than either the Hartree-Fock or B3LYP calculations.

### 8.2.3 Conclusions

It is difficult to draw any definitive conclusions about the effect of deuteration on the  $^{15}\text{N}$  CSA from these calculations, since the results of the various calculations did not agree even inasmuch as regarding the direction of the trends. The one

conclusion that can be drawn is that the effect of deuteration on the  $^{15}\text{N}$  CSA is expected to be small with respect to the  $^{15}\text{N}$  CSA itself; calculated values of  $\Delta\Delta\sigma^{15\text{N}}(^1\text{H}\rightarrow^2\text{H})$  ranged from -1.1 to 7.8 ppm, and in all cases was smaller in absolute value than 8 ppm. Even in the MP2 calculation, which predicted a largest change in the CSA due to deuteration (7.8 ppm), this effect was only at the level of ~5% of the CSA in the  $^{15}\text{N}\{-^1\text{H}\}$  system. This is comparable to the level of experimental uncertainty in our measurement of the CSA from the field dependence of relaxation data.

### **8.3 Pulse Sequences, Spectra, and Rates**

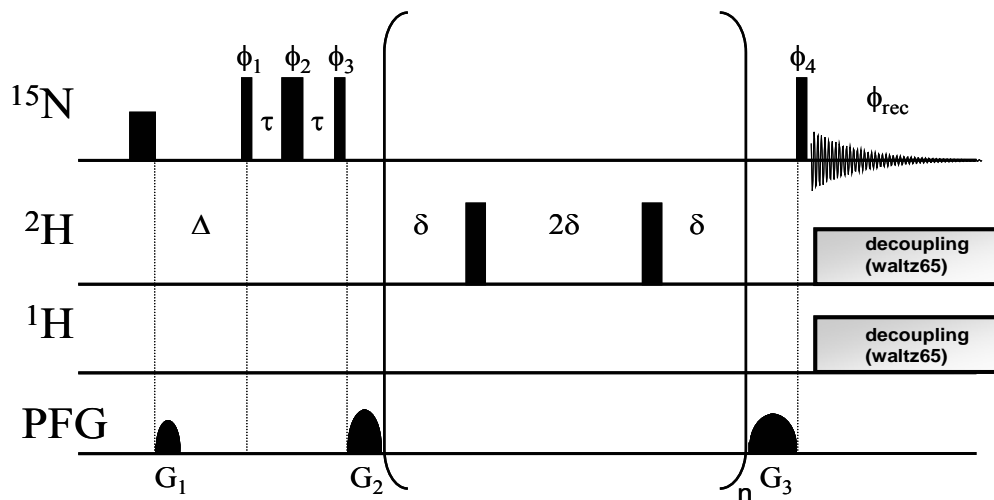
#### **8.3.1 Pulse Sequences for Measurement of $^{15}\text{N}$ $R_1$ and $^{15}\text{N}$ $R_2$ using Direct Nitrogen Detection.**

Pulse sequences for 1D  $^{15}\text{N}$ -detected  $R_1$  and  $R_2$  measurements are shown in Figure 8.3.1 (a) and (b). The relaxation delays for the  $R_1$  experiments were 11.15, 105.84, 209.13, 295.21, 398.51 (x2), 605.1, 794.48 and 1001.07 ms and for the  $R_2$  experiments 3.74, 103.10, 198.78 (x2), 265.02, 323.90, 397.50 and 449.02 ms. The pulse repetition delays  $\delta$  were of 2 ms and 200  $\mu\text{s}$  in the  $R_1$  and  $R_2$  experiments, respectively. The delay  $\tau$  was 10  $\mu\text{s}$ . The phases were:  $\phi_1=(y,-y)$ ,  $\phi_2=(8x,8(-x))$ ,  $\phi_3=(2(-y),2y)$ ,  $\phi_4=(4x,4(-x))$  and  $\phi_{\text{rec}}=(x,2(-x),x,-x,2x,-x)$  in the  $R_1$  experiment and  $\phi_1=(y,-y)$ ,  $\phi_2=(2x,2(-x))$  and  $\phi_{\text{rec}}=(x,2(-x),x)$  in the  $R_2$  experiment (the latter phase cycle is based on a recently published modification of the CPMG experiment <sup>26</sup>). The number of scans was from 4k to 8k, depending on the total relaxation period. Waltz65 decoupling was applied during the acquisition on the  $^1\text{H}$  and  $^2\text{H}$  channels; this

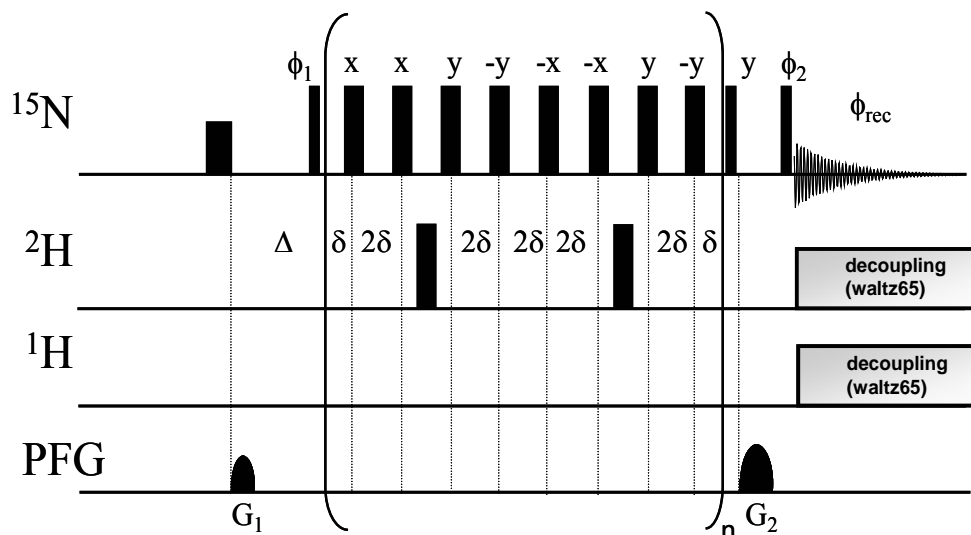
consists of an MLEV4 supercycle of the basic waltz16 element (waltz64) with an additional  $90^\circ$  pulse at the end of the supercycle for increased performance with respect to decoupling sidebands. The acquisition time was 300 ms.  $^{15}\text{N}$  pulses were applied at 16.6 kHz, while the  $^2\text{H}$  decoupling was at a power level of 880 Hz. A low-power  $90^\circ$  pulse followed by a gradient was applied at the beginning of the sequences in order to defocus the  $^{15}\text{N}$  magnetization prior to the relaxation period. The  $^2\text{H}$  spectrometer lock was kept on during the relaxation period  $\Delta$  (2.9 s) and switched off just prior to the first rf pulse.

The measurements were performed on a 5.8 mM sample of protein G (GB3)  $^{152}$  in  $\text{D}_2\text{O}$  at 11.7 T and  $24^\circ\text{C}$ , on a 5mm Z-Gradient P/C/N-H/D QNP CryoProbe using a cryogenically cooled preamplifier for all nuclei. The cooled  $^{15}\text{N}$  preamplifier in the QNP probe offers a 4-fold gain in sensitivity compared to conventional broadband and QNP probeheads and approximately similar sensitivity increase with respect to the TXI or TCI cryoprobes. Our data indicate that the proposed measurements are still feasible using broadband and cryo-TXI probes, on fairly concentrated protein samples.

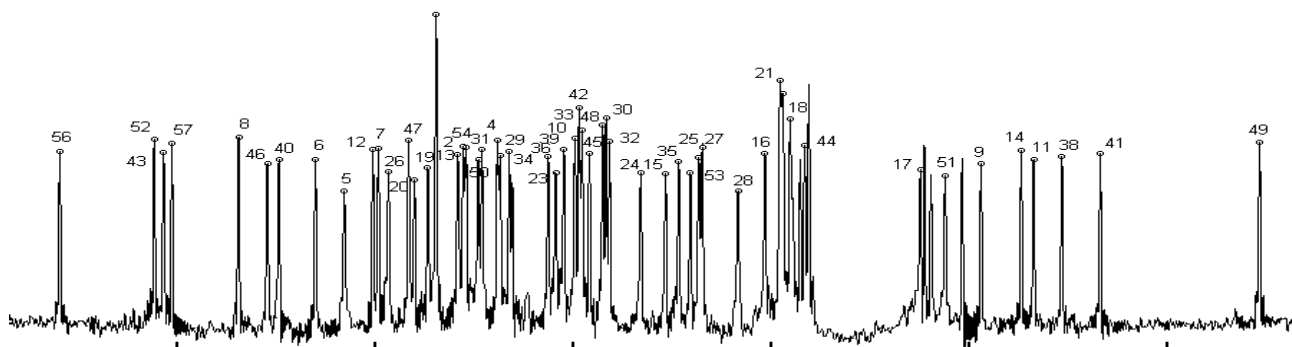
**Figure 8.3.1.** Pulse sequences for 1D  $^{15}\text{N}$ -detected  $R_1$  measurement. The phases cycle is:  $\phi_1=(y,-y)$ ,  $\phi_2=(8x,8(-x))$ ,  $\phi_3=(2(-y),2y)$ ,  $\phi_4=(4x,4(-x))$  and  $\phi_{\text{rec}}=(x,2(-x),x,-x,2x,-x)$ . Sine-bell shaped gradients with a duration of 1 ms and ratios of intensities of  $G_1:G_2:G_3=-5:2:1.4$ . The number of scans was from 4k to 8k, depending on the total relaxation period. Waltz65 decoupling was applied during the acquisition on the  $^1\text{H}$  and  $^2\text{H}$  channels; this consists of an MLEV4 supercycle of the basic waltz16 element (waltz64) with an additional  $90^\circ$  pulse at the end of the supercycle for increased performance with respect to decoupling sidebands. The acquisition time was 300 ms.  $^{15}\text{N}$  pulses were applied at 16.6 kHz, while the  $^2\text{H}$  decoupling was at a power level of 880 Hz.



**Figure 8.3.2.** Pulse sequences for 1D  $^{15}\text{N}$ -detected  $R_2$  measurement. The phases cycle is:  $\phi_1=(y,-y)$ ,  $\phi_2=(2x,2(-x))$  and  $\phi_{\text{rec}}=(x,2(-x),x)$ . Sine-bell shaped gradients with a duration of 1 ms and ratios of intensities of  $G_1:G_2=-5:2$ . The number of scans was from 4k to 8k, depending on the total relaxation period. See caption to Figure 8.2.2 for further experimental parameters.

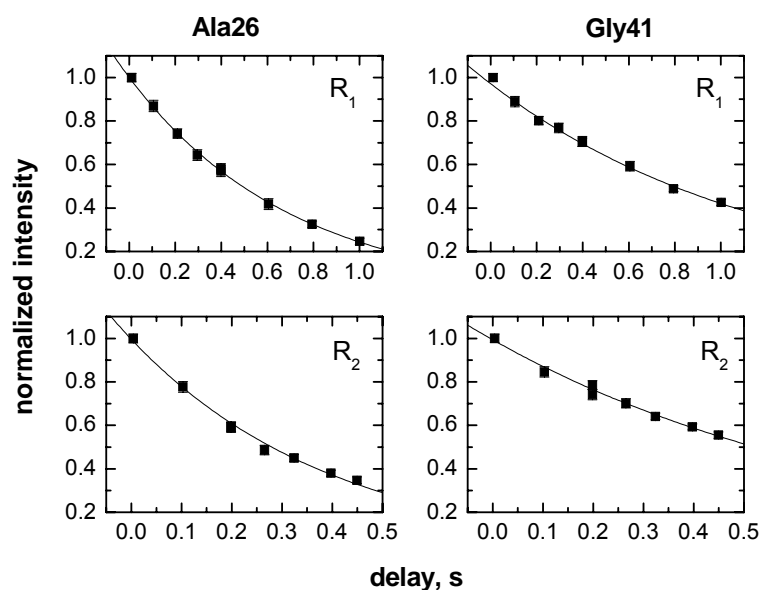


**Figure 8.3.3.**  $^{15}\text{N}$ -detected spectrum of GB3 (pH 5.5, 24°C), recorded at 11.7 T with 4096 scans and a relaxation delay of 2.9 s (total time 4h). The residue numbers are indicated at the frequency position of the corresponding  $^{15}\text{N}$  resonance. All the assigned signals originating from backbone amides that could be resolved are labeled with the residue number, the indole signal of Trp43 is labeled 57. Circles without labels correspond to signals in overlap.



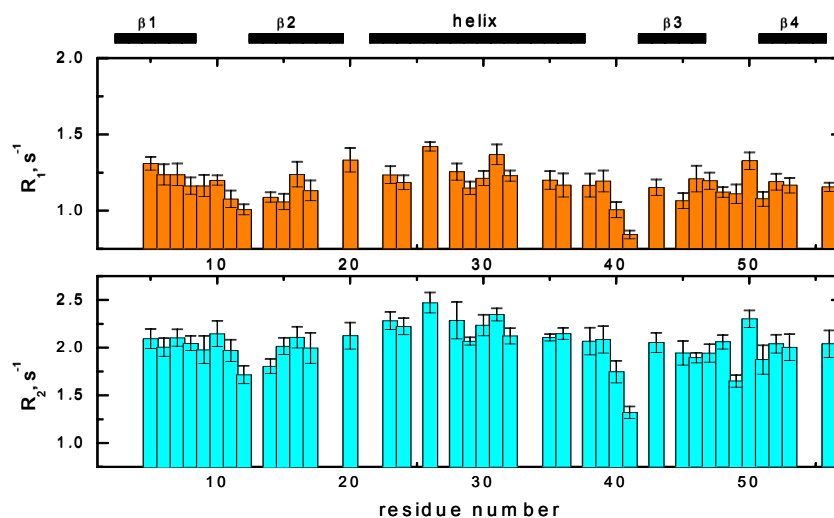
### 8.3.2 Measured Rates

**Figure 8.3.4.** Representative fitting curves for  $R_1$  and  $R_2$  measurements are shown for Glu24 and Lys50, as indicated. Peak intensities in a series of 1D spectra recorded with different relaxation delays were fit to a mono-exponential decay.





**Figure 8.3.5.** Longitudinal ( $R_1$ ) (cyan) and transverse ( $R_2$ ) (orange)  $^{15}\text{N}$  relaxation rates in GB3 in  $\text{D}_2\text{O}$ , determined using  $^{15}\text{N}$  direct detection at 11.5 T (500 MHz  $^1\text{H}$  frequency). The cartoon at the top indicates the location of the secondary structure elements.



The  $^{15}\text{N}$  relaxation rates  $R_2$  and  $R_1$  were measured using the pulse sequences shown in Figures 8.3.1 and 8.3.2. Representative fitting curves are shown in Figure 8.3.3. The  $^{15}\text{N}$  relaxation data are depicted as bars in Fig. 8.3.4 as a function of residue number. The profile here is similar to that of the relaxation data in  $\text{H}_2\text{O}$  (cf. Chapter 5 Figure 5.3.1), in that the transverse relaxation rates in the  $\alpha$ -helix are slightly higher than those measured in the rest of the protein, although the difference is not as striking as for the  $\text{H}_2\text{O}$  data. This suggests that the elevated relaxation rates of amide nitrogens in the helix are not due to higher CSA values, otherwise the rates in this region would be more elevated in the data recorded in  $\text{D}_2\text{O}$  compared to data in  $\text{H}_2\text{O}$ . As was shown in Chapter 5, the profile of relaxation rates in GB3 is mainly determined by the overall rotational diffusion properties of the molecule; the elevated  $R_2$ s in the  $\alpha$ -helix are the result of the unique principal axis of the diffusion tensor of GB3 being approximately parallel to the helix axis (hence to NH vectors in the helix)

and approximately perpendicular to NH vectors in the beta sheet. This effect is less pronounced in D<sub>2</sub>O, because the CSA tensors – the main contributors to <sup>15</sup>N relaxation – in the beta-sheet residues are tilted from the NH vectors, and are therefore more aligned with the main diffusion axis (a 17° tilt of the CSA in the peptide plane corresponds to an average 10° tilt *towards* the diffusion tensor for beta sheet residues). The CSA tensors in the alpha-helix are correspondingly less well aligned with the helical axis (and therefore with the diffusion tensor axis) than their NH vectors, due to this same tilt, the combined effect being that there is less of an elevation of transverse relaxation rates in the helix with respect to the transverse relaxation rates in the sheets. The low values of both relaxation rates in the β1/β2 and α/β3 loops are due to increased mobility in these regions.

#### **8.4 Diffusion Tensor of GB3 in D<sub>2</sub>O Solution Derived from <sup>15</sup>N Direct Detection Relaxation Rates: Proof of Principle**

In order to validate the relaxation data obtained by <sup>15</sup>N direct detection, we used them to determine the overall rotational diffusion tensor of GB3, for comparison with this tensor derived in H<sub>2</sub>O.

##### **8.4.1 Subtraction of Contributions to the Relaxation Rates Based On Model-Free Parameterization**

Recall that the method for fitting the diffusion tensor of a molecule uses a ratio of “reduced” relaxation rates (Eq. 5.2.2), usually obtained by subtracting the components of the spectral density involving the hydrogen frequency (in the NH

pair). This is done so that the ratio,  $\rho$ , is independent, to a first approximation, of the  $^{15}\text{N}$  CSA values and of the order parameters describing local backbone motion. To adapt this method to ND systems, the spectral densities  $J(\omega)$  at frequencies involving deuterium frequency ( $\omega_D$ ,  $\omega_D \pm \omega_N$ ) were subtracted from the measured  $R_1$  and  $R_2$  values as follows:

$$R_1' = R_1 - \frac{8}{3} d^2 [6J(\omega_D + \omega_N) + J(\omega_D - \omega_N)], \quad (8.4.1)$$

$$R_2' = R_2 - \frac{4}{3} d^2 [6J(\omega_D) + 6J(\omega_D + \omega_N) + J(\omega_D - \omega_N)], \quad (8.4.2)$$

where  $d = -\frac{1}{2} \frac{\mu_0}{4\pi} \frac{\hbar \gamma_N \gamma_D}{r_{ND}^3}$  is the dipolar coupling constant in the ND pair. The

spectral densities were calculated using the model-free assumption<sup>27</sup>. As shown in<sup>25</sup>, the model-free form of the spectral density is applicable to ND bond dynamics in deuterated amides. Given the tensor in  $\text{H}_2\text{O}$  was shown to be axially symmetric to a good approximation, axial symmetry was assumed for the overall rotational diffusion tensor here. The ND bond length was set to 1.02 Å. The subtracted  $J(\omega)$  components (Eqs. 8.4.1-8.4.2) and the  $R_2'/R_1'$  ratio are independent of  $^{15}\text{N}$  CSA, therefore the diffusion tensor can be derived without making any assumption regarding the CSA values<sup>120</sup>. The calculation of the  $J(\omega)$  components to be subtracted assumed the diffusion tensor determined from  $^{15}\text{N}$  relaxation data for GB3 in  $\text{H}_2\text{O}$  at 14.1 T (Chapter 5, and see below). In addition to this partial subtraction of the spectral densities associated with N-D dipolar interaction, a second calculation was performed in which the entire dipolar contribution to the relaxation rates was subtracted, as follows:

$$R_1' = R_1 - 3d^2 J(\omega_N) - \frac{8}{3} d^2 [6J(\omega_D + \omega_N) + J(\omega_D - \omega_N)], \quad (8.4.3)$$

$$R_2' = R_2 - \frac{1}{2} d^2 [4J(0) + 3J(\omega_N)] - \frac{4}{3} d^2 [6J(\omega_D) + 6J(\omega_D + \omega_N) + J(\omega_D - \omega_N)] \quad (8.4.4)$$

### 8.4.2 Diffusion Tensor of GB3: Proof of Principle

Since a large fraction of the  $^{15}\text{N}$  relaxation in ND pairs is due to CSA modulation (see Table 8.1.2), and the orientations of the principal components of the CSA tensor are unknown *a priori*, it is unclear whether the derivation of the overall diffusion tensor from these data using the same methodology applied to relaxation rates in NH pairs will result in reasonable values. The problem is as follows: the method for derivation of the overall diffusion tensor (e.g. using the ROTDIF algorithm) uses the orientations of the NH bond vectors from the crystal structure of the protein as representative of the directions of the relaxation-active interactions. While this is probably a good approximation in NH systems where the dipolar interaction is the major contribution to relaxation, in ND pairs where the CSA interaction is more important (see Table 8.1.2), the orientation of the relaxation-active interaction may (1) not be vectorial (if the  $^{15}\text{N}$  CSA is not axially symmetric), and (2) provided the vector representation is applicable, the orientation of this “CSA vector” may not be along the ND bond. However, at 500 MHz there is still a significant contribution to ND relaxation rates from the dipolar interaction (see Table 8.1.2), there is some evidence that the  $^{15}\text{N}$  CSA is approximately axially symmetric (see Table 1.3.1), and that the angle between this symmetry axis and the bond vector is

small ( $\sim 20^\circ$ - $30^\circ$ , see Figure 7.4.2b). Given all of this, it is interesting to see if the conventional methods for derivation of the rotational diffusion tensor give results which are in agreement with those derived on the GB3 sample in  $\text{H}_2\text{O}$ .

The anisotropies and orientations of the derived diffusion tensor are shown in Table 8.4.1 in comparison with those determined from  $^{15}\text{N}$  relaxation rates measured in  $\text{H}_2\text{O}$  using conventional methods. The observed slower tumbling of GB3 in  $\text{D}_2\text{O}$  ( $\tau_c = 4.2$ - $4.3$  ns versus  $3.3$  ns in  $\text{H}_2\text{O}$ ) is consistent with a 1.25-fold higher viscosity of  $\text{D}_2\text{O}$  solvents<sup>177</sup>. This tumbling time also agrees with the value of  $4.55 \pm 0.24$  ns obtained from  $^{15}\text{N}$  relaxation data measured for the residual NH groups (approximately 2-3%) in this GB3 sample in  $\text{D}_2\text{O}$  using conventional INEPT-based 2D methods. Furthermore this tensor determination procedure is self-consistent in that the output  $\tau_c$  value from the ROTDIF calculation agrees well with the input value used for the subtraction of dipolar contributions.

**Table 8.4.1. Characteristics of the overall rotational diffusion tensor of GB3 determined in H<sub>2</sub>O and D<sub>2</sub>O solvent**

| Spin pair   | $J(\omega)$ components subtracted                          | $\beta^a$       | $\xi^a$        | $\tau_c^b$ | $D_{  }/D_{\perp}^b$ | $\Phi^c$ | $\Theta^c$ |
|---|--|-----------------|----------------|------------|----------------------|----------|------------|
| <sup>15</sup> N- <sup>1</sup> H<br>(sample in H <sub>2</sub> O) | $J(\omega_H), J(\omega_H \pm \omega_N)$                    | 0 <sup>d</sup>  | 0 <sup>d</sup> | 3.3 (0.1)  | 1.4 (0.1)            | 92 (6)   | 72 (7)     |
|   |  | -3 <sup>e</sup> | 0 <sup>e</sup> | 3.3 (0.1)  | 1.4 (0.1)            | 95 (6)   | 68 (7)     |
| <sup>15</sup> N- <sup>2</sup> H<br>(sample in D <sub>2</sub> O) | $J(\omega_D), J(\omega_D \pm \omega_N)$                    | 0 <sup>d</sup>  | 0 <sup>d</sup> | 4.2 (0.4)  | 1.3 (0.2)            | 103 (28) | 68 (23)    |
|   |  | 17 <sup>e</sup> | 2 <sup>e</sup> | 4.2 (0.4)  | 1.3 (0.2)            | 96 (26)  | 75 (21)    |
|   | $J(0), J(\omega_N), J(\omega_D), J(\omega_D \pm \omega_N)$ | 0 <sup>d</sup>  | 0 <sup>d</sup> | 4.3 (0.5)  | 1.3 (0.3)            | 105 (34) | 66 (29)    |
|   |  | 24 <sup>e</sup> | 3 <sup>e</sup> | 4.3 (0.4)  | 1.3 (0.3)            | 98 (26)  | 77 (23)    |

Numbers in the parentheses represent standard deviations.

<sup>a</sup> The angles  $\beta$  and  $\xi$  (in degrees) correspond to a tilt of the interaction axis away from the NH bond (see text).

<sup>b</sup> The overall correlation time  $\tau_c = 1/[2(D_{||}+2D_{\perp})]$  (in nanoseconds) and the anisotropy,  $D_{||}/D_{\perp}$ , of the diffusion tensor.

<sup>c</sup> The angles  $\Phi$  and  $\Theta$  (in degrees) determine the orientation of unique axis of the rotational diffusion tensor with respect to the protein coordinate frame.

<sup>d</sup> The angles  $\beta$  and  $\xi$  were fixed at 0, i.e. the interaction axis was assumed to be in the direction of the N-H bond.

<sup>e</sup> The angles  $\beta$  and  $\xi$  were adjusted as a result of a grid search minimizing the residuals of fit.

### 8.4.3 Relative Orientation of the CSA and Dipolar Relaxation Mechanisms

We then used our  $^{15}\text{N}$  relaxation data in  $^{15}\text{N}\text{-}\{^2\text{H}\}$  systems to explore the orientation of the CSA tensor with respect to the peptide plane. The determination of the diffusion tensor of a protein from  $^{15}\text{N}$  relaxation data uses the dependence of the relaxation rates on the angles between the principal axes of the tensor and the symmetry axis of the relaxation-active terms in the spin Hamiltonian. In  $^{15}\text{N}\text{-}\{^1\text{H}\}$  systems the orientation dependence of the  $^{15}\text{N}\text{-}^1\text{H}$  dipolar interaction is the primary contribution to the time-varying magnetic field at the site of the  $^{15}\text{N}$  nucleus, and thus the primary contribution to the orientational dependence of  $^{15}\text{N}$  relaxation rates comes from the angle between the NH bond and the axis of the diffusion tensor. In  $^{15}\text{N}\text{-}\{^2\text{H}\}$  spin systems, the CSA contribution is the larger term and therefore the orientation of the  $^{15}\text{N}$  CSA tensor becomes the determining factor for the orientation dependence of the  $^{15}\text{N}$  relaxation rates. Because the chemical shift is sensitive to the local distribution of electron density, the  $^{15}\text{N}$  CSA tensor's orientation is in general different from that of the NH-bond vector. The anisotropy of local backbone motions could also contribute to the difference between average orientations of the CSA and dipolar tensors. To account for the possible difference in the orientations of the relaxation-active interaction and the NH vector, we allowed the modeled symmetry axis of the chemical shift to deviate from the NH bond in both in-plane and out-of-plane directions using two degrees of freedom described by the angles  $\beta$  and  $\xi$ . The angle  $\beta$  measured the deviation of the interaction frame from the NH orientation in the peptide plane, with the positive direction towards the carbonyl nucleus of the

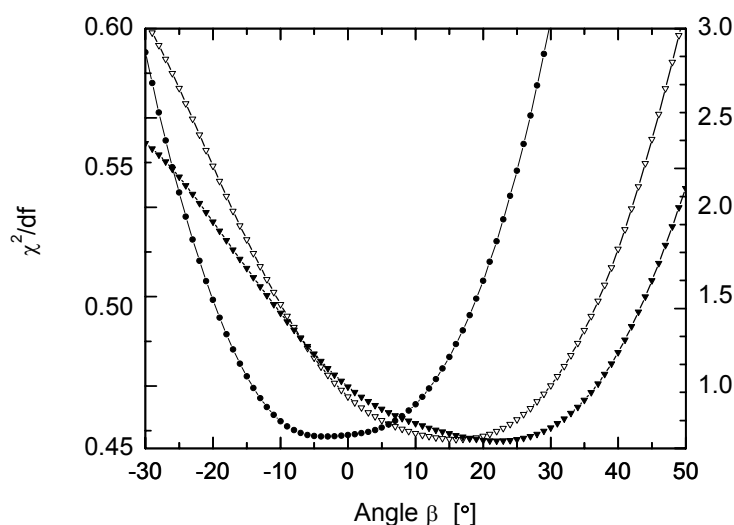
previous residue (i.e. here  $\beta$  measures  $\beta_z$  in Fig.8.2.2). The angle  $\xi$  measured the deviation away from the peptide plane, with the positive sign corresponding to a clockwise rotation about the N-H bond (when looking from H to N). The tilt of the interaction axis was assumed to be the same for all residues in the protein. The ROTDIF analysis was performed on a  $\{\beta, \xi\}$  grid, and the resulting values of the target function obtained from all grid points were compared, in order to find the orientation of the axis that minimizes the difference between the experimental and back-calculated values of  $\rho$ .

Using reduced relaxation rates (Eq. 8.4.2), the optimal orientation of the axis of interaction was found to be approximately in the peptide plane and tilted by  $17^\circ$  from the NH-bond-vector towards the carbonyl atom. When the entire dipolar contribution was subtracted from the relaxation rates, this fit resulted in a further increase, to  $24^\circ$ , in the tilt angle  $\beta$  in the same direction of rotation. Given the relatively broad minimum of the target function versus angle  $\beta$ , these numbers are in good agreement with one another. Although the reduction in  $\chi^2$  compared to zero-tilt model is not dramatic, likely reflecting the limited precision of the data, the statistical F-test gave confidence levels of 83 and 85%, respectively, for the two levels of subtraction. The out-of-plane tilt angle  $\xi$  was small ( $2^\circ$ - $3^\circ$ ) and statistically insignificant. The slight increase in the tilt angle when the dipolar contribution to relaxation was completely subtracted agrees with the idea that the observed behavior is the result of an interplay between the orientations of the dipolar and CSA interactions. In a control analysis of  $^{15}\text{N}$  relaxation data for GB3 in  $\text{H}_2\text{O}$ , the



minimum was found within  $3^\circ$  of the NH vector orientation, as expected (see Figure 8.4.1).

**Figure 8.4.1. Variation in the residuals of the fit (from ROTDIF analysis), as the main interaction frame is tilted away from the direction of the NH vector, while remaining in the peptide plane (angle  $\xi = 0$ ).** Shown is the value of the target function per degree of freedom ( $\chi^2/\text{df}$ ) from a least-square fit of experimental data assuming axial symmetry of the overall rotational diffusion tensor. The curves correspond to  $\text{D}_2\text{O}$  rates with subtraction of the dipolar contributions at frequencies containing combinations of the deuterium Larmor frequency (Eqs.8.4.2, open triangles) and at all frequencies (Eqs.8.4.3, solid triangles) and the  $\text{H}_2\text{O}$  rates (solid circles)



## 8.5 Conclusions

Here we have used heteronuclear-detected experiments to measure  $^{15}\text{N}$  relaxation rates in GB3 in  $\text{D}_2\text{O}$ , where the dipolar contribution is lessened by replacing the dipolar-coupled partner ( $^1\text{H}$ ) with a deuteron. We have shown that these rates have increased relative sensitivity to the  $^{15}\text{N}$  CSA compared to relaxation rates in  $^{15}\text{N}$ - $^1\text{H}$  systems. This additional sensitivity allowed us to estimate the average of the angle between the  $^{15}\text{N}$  CSA and the  $^{15}\text{N}$ - $^2\text{H}$  bond from the orientation dependence of a ratio of  $^{15}\text{N}$  relaxation rates. This angle was found to be between  $17\text{-}24^\circ$  to  $80\%$

confidence in agreement with previous estimations from combinations of auto and cross-correlation rates at several fields (Chapter 7, Figure 7.4.2c). In order to determine that the CSA of  $^{15}\text{N}$  in  $^{15}\text{N}\text{-}\{^2\text{H}\}$  systems was similar in magnitude to the  $^{15}\text{N}$  CSA in  $^{15}\text{N}\text{-}\{^1\text{H}\}$  systems, we performed quantum mechanical *ab initio* calculations at the B3LYP6-311+G(2d,p) level on the N-methylacetamide molecule.  $\Delta\Delta\sigma^{^{15}\text{N}}(^1\text{H}\rightarrow^2\text{H})$  from these studies were small compared to the  $^{15}\text{N}$  CSA, and on order of the expected error for CSA determination using  $^{15}\text{N}$  relaxation rates.

## Chapter 9: Summary and Concluding Remarks

### **9.1 Summary of Results**

New direct methods for measuring the  $^{15}\text{N}$  CSA/dipolar cross-correlation rates (CCRs) in proteins were presented and shown to be in good agreement with indirect methods when an arbitrary scaling factor (here observed to be between 4-7%) is applied to the cross-correlation rates measured by the indirect methods. It was shown here that  $^{15}\text{N}$  CSA/dipolar CCRs can be measured directly from the relative amplitudes of the up- and down-field  $^{15}\text{N}$  signals in a  $^1\text{H}$ -coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum. The obvious advantage of this approach is that both signals are observed in the same spectrum and, therefore, no ambiguity associated with correction factors is involved. The application of this method to biological macromolecules, however, is complicated by signal overlap in the coupled 2D spectra, which may be particularly severe in the case of  $^1\text{H}$ -coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for macromolecules greater than 10 kDa. With this in mind, two direct methods were proposed for  $^{15}\text{N}$ /CSA dipolar CCR measurement (the IPAP and  $\text{S}^3\text{E}$  methods) based on spectral simplification schemes. It was shown that no scaling factor is needed for these direct methods, since the CCRs from the direct methods are in agreement with the CCRs obtained directly from the time evolution of the ratio of the components of the  $^{15}\text{N}$  doublet in a coupled HSQC-type spectrum.

It was shown that the IPAP scheme simplifies coupled  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra without causing deviations in cross-correlation rates from those measured with the simple IP scheme. Both CCRs measured using the IP only sequence and the IPAP

scheme agree with the indirect A/B method when the A/B method is scaled by a correction factor. There is no such need for a scaling factor between the IP and IPAP derived CCR.  $S^3E$  spin-state selection of the individual (up- or down-field) component of the nitrogen doublet before the relaxation delay also alleviates problems due to spectral overlap. This experiment is, in principle, even better than that of the IPAP scheme, which might introduce very small errors in restored-peak intensities of overlapping peaks due to site-specific differences in relaxation properties of individual amides. CCRs measured using the sequence containing the  $S^3E$  spin-state selection element agree with IPAP, IP, and scaled A/B method CCRs. There is no need for a scaling factor between the  $S^3E$  and IP derived CCRs.

The overall rotational diffusion tensor and correlation time of the GB3 domain were determined from  $^{15}N$  relaxation rates ( $R_1$ ,  $R_2$ ,  $^{15}N\{^1H\}$  NOE) at 600 MHz. It was found that to a good approximation, GB3 can be modeled as a prolate axially symmetric (symmetric-top) rotor with the ratio of the rate of reorientation about the fast axis to a perpendicular axis,  $D_{\parallel} / D_{\perp}$  of 1.37 and an overall correlation time,  $\tau_c$  of 3.34 ns. The improvement in the fit of the relaxation data using the fully anisotropic diffusion tensor ( $D_z / D_y = 1.36$  and  $D_y / D_x = 1.11$ ) was not found have statistical significance compared to the fit using the axially symmetric diffusion tensor. These experimentally determined parameters describing the overall diffusion of the GB3 domain were shown to be in general agreement with the predictions of theoretical hydrodynamic models. Furthermore, this diffusion tensor derived from data at 600 MHz was found to be in excellent agreement with the GB3 diffusion tensor derived from measurements at additional spectrometer fields, 400-800 MHz,

with a globally-fit diffusion tensor using all the data combined, and with the diffusion tensor derived from cross-correlation rates at fields 400-600 MHz.

An analysis of Lipari-Szabo model-free parameters describing the motion of NH bonds in the polypeptide backbone of the GB3 domain from  $^{15}\text{N}$  relaxation data shows that the isotropic and anisotropic models of the overall tumbling result in markedly different pictures of local motion; the main difference is in the interpretation of the elevated  $R_2$  values in the  $\alpha$ -helix: the isotropic model results in conformational exchange throughout the  $\alpha$ -helix, whereas no exchange is predicted by anisotropic models that place the longitudinal axis of diffusion tensor almost parallel to the helix axis. Both axially symmetric and fully anisotropic models for the overall motion fit the experimental data significantly better than does a model with an isotropic diffusion tensor. Based on statistical F-tests, we conclude that the overall rotational diffusion of GB3 is best modeled as axially symmetric, and that the fully anisotropic description of the diffusion tensor is not statistically warranted.

Since the results of the Lipari-Szabo analysis using the anisotropic models of overall diffusion could be biased (see Chapter 4), additional, model-independent methods for identification of exchange motions which do not rely on knowledge of protein structure or assumptions about its dynamics were required in order to unambiguously distinguish the correct picture of local motion of peptide planes. Three such methods were applied to differentiate between the effects of conformational exchange and rotational anisotropy: a comparison of the CSA/dipolar cross-correlation rates ( $\eta_{xy}$ ,  $\eta_z$ ) with relaxation rates ( $R_2$ ,  $R_1$ ), the estimation of  $R_{ex}$  terms from  $^{15}\text{N}$  relaxation data at two fields, and relaxation-compensated CPMG

measurement of exchange from experiments with different values of  $\nu_{\text{CPMG}}$  to identify exchange on longer timescales (1-8 ms). These methods are all sensitive to conformational exchange, and do not require knowledge of the protein structure or any assumptions about spectral density functions, and therefore can be used to either identify potential sites for conformational exchange (or verify the absence thereof). These analyses provide no indication of conformational exchange in the helix, consistent with predictions of the Lipari-Szabo analysis using the axially symmetric or fully anisotropic overall diffusion tensor.

A comprehensive study of the  $^{15}\text{N}$  chemical shielding anisotropy in the GB3 domain was presented based on a combination of  $^{15}\text{N}$  relaxation and  $^{15}\text{N}$  CSA/dipolar cross-correlation measurements at five static magnetic fields. The analysis was performed using various combinations of the experimental data and using model-independent approaches as well as methods based on Lipari-Szabo approximation. The results indicate significant site-to-site variations in the principal values and the orientation of the  $^{15}\text{N}$  CSA, similar to those observed earlier in ubiquitin<sup>51,52</sup>. Our estimates of the true variability in the  $^{15}\text{N}$  CSA in GB3 depend to some degree upon which method for determining the CSA was used and which subset of residues is considered. These estimates range from 10.2 ppm (for 33 residues that pass the  $\chi^2/\text{df}$  cutoff from the  $R/\eta$  method) to 21.4 ppm for all 47 residues from the  $2R_2-R_1$  method. Although this range of values could be a result of limited statistics, all of these estimates are still larger than the derived variability in the  $^{15}\text{N}$  CSA from studies of ribonuclease H<sup>53</sup> or recently of ubiquitin using a subset of the methods used here<sup>54</sup>. The true mean CSA values range from  $-173.9$  ppm ( $2R_2-R_1$ ) to  $-177.2$  ppm ( $R/\eta$ ).

Our data show that using the site-specific values of the  $^{15}\text{N}$  chemical shielding anisotropy obtained here significantly improves the agreement between LS order parameters measured at different fields and allows simultaneous fit of the  $^{15}\text{N}$  relaxation data at five fields to LS spectral densities. These findings emphasize the necessity of taking into account the variability of the  $^{15}\text{N}$  chemical shielding tensor for accurate analysis of protein dynamics from  $^{15}\text{N}$  relaxation measurements. This can be achieved by including the  $^{15}\text{N}$  CSA as an additional fitting parameter in the LS analysis of multiple-field data, provided the sample temperature and other experimental conditions are the same at all fields/spectrometers. These analyses also show that the Lipari-Szabo form of the spectral density provides a satisfactory approximation for the experimental spectral densities obtained using the reduced spectral density approach.

Novel methods for measurement of  $^{15}\text{N}$  relaxation rates in a protein in  $\text{D}_2\text{O}$  using direct  $^{15}\text{N}$  detection were presented and demonstrated on the GB3 domain. By sampling the spectral density function at the frequencies  $\omega_D$  and  $\omega_D \pm \omega_N$  these measurements provide potentially useful information about protein motions in the nanosecond time range not available from the conventional measurements in  $\text{NH}$  systems. In addition, the proposed direct  $^{15}\text{N}$ -detection experiments offer increased sensitivity to  $^{15}\text{N}$  CSA values and could provide a useful tool for accurate measurements of these parameters in proteins. Quantum mechanical *ab initio* calculations at the B3LYP6-311+G(2d,p) level were conducted on the N-methylacetamide molecule in the gas phase to explore the sensitivity of the  $^{15}\text{N}$  chemical shielding anisotropy to isotopic substitution of the directly bound hydron.

$\Delta\Delta\sigma^{15N}(^1H\rightarrow^2H)$  from these studies ranged from -1.1 to 7.8 ppm, and in all calculations was smaller in absolute value than 8 ppm, and therefore on order of the expected error for CSA determination using direct detected  $^{15}N$  relaxation rates.

## **9.2 Scope for Future Studies**

$^{15}N$  NMR spin relaxation rates contain a wealth of information about chemistry, structure, and dynamics in biomolecules. However, there is not yet a generalized “best approach” to deconvolution of the individual contributions to these rates of a multitude of interaction and motional parameters (e.g. dipolar coupling constants, chemical shift tensors, characteristics of the overall and local motions). The development of such a generalized approach was not the goal of this work, but a necessary step towards such an approach is an exploration of the boundaries of applicability of current methods for dynamics analysis. Specifically in question here were the degree of overall diffusional anisotropy that can safely be neglected in determination of the local backbone peptide plane dynamics, and the effect on these dynamics of variability of the  $^{15}N$  CSA from residue to residue within proteins.

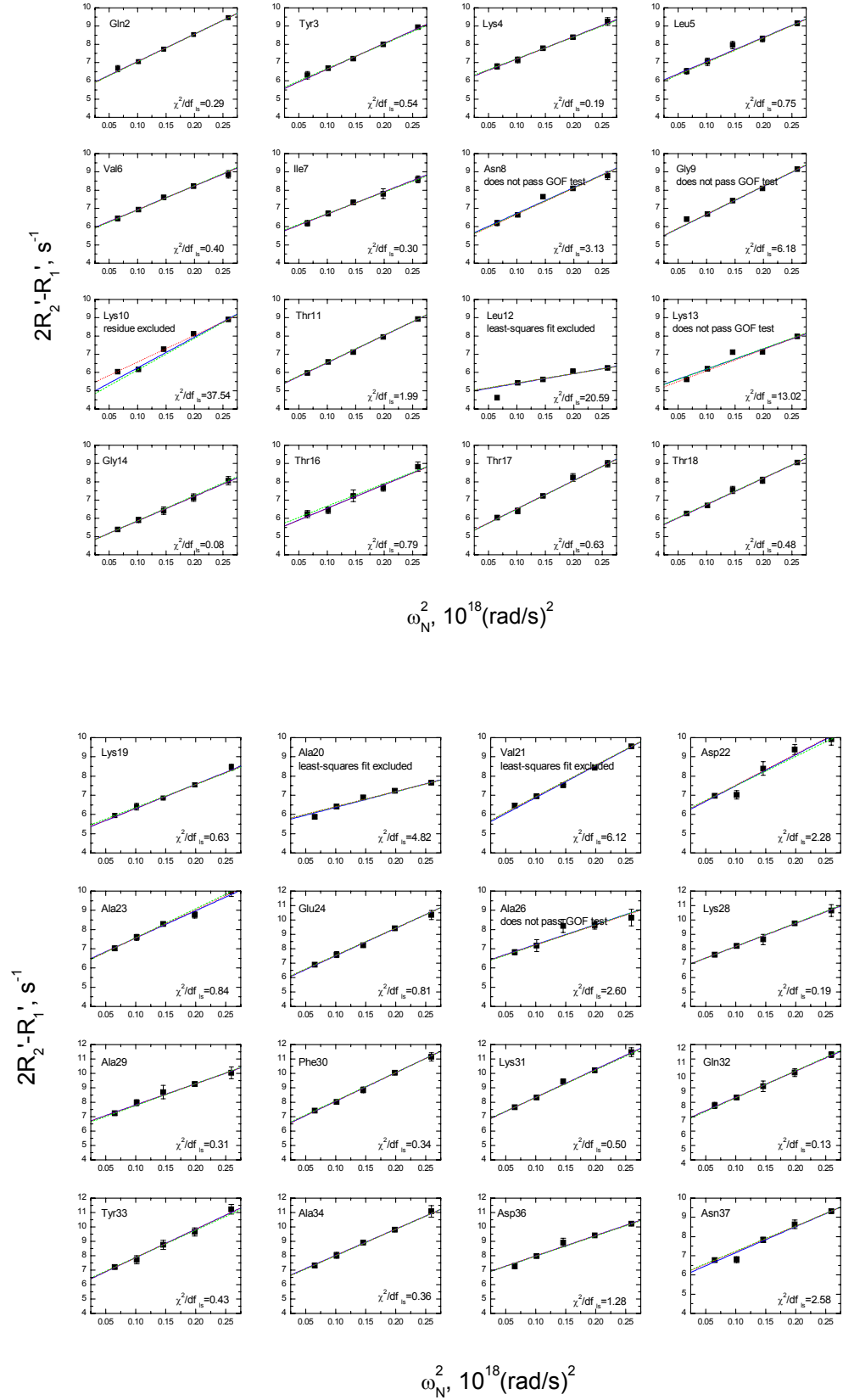
It was found that for the GB3 domain, a Lipari-Szabo analysis using an isotropic model of the overall rotational diffusion resulted in spurious conformational exchange motions and an underestimation of order parameters for all residues in the  $\alpha$ -helix. When an anisotropic model of overall diffusion is assumed, it becomes clear that there is no significant conformational exchange motion in the GB3 domain with the possible exception of residue Val39. However, there seems to be no significant difference between the Lipari-Szabo parameters derived using the axially symmetric and fully anisotropic tensors. The overall rotational diffusion of GB3 is axially

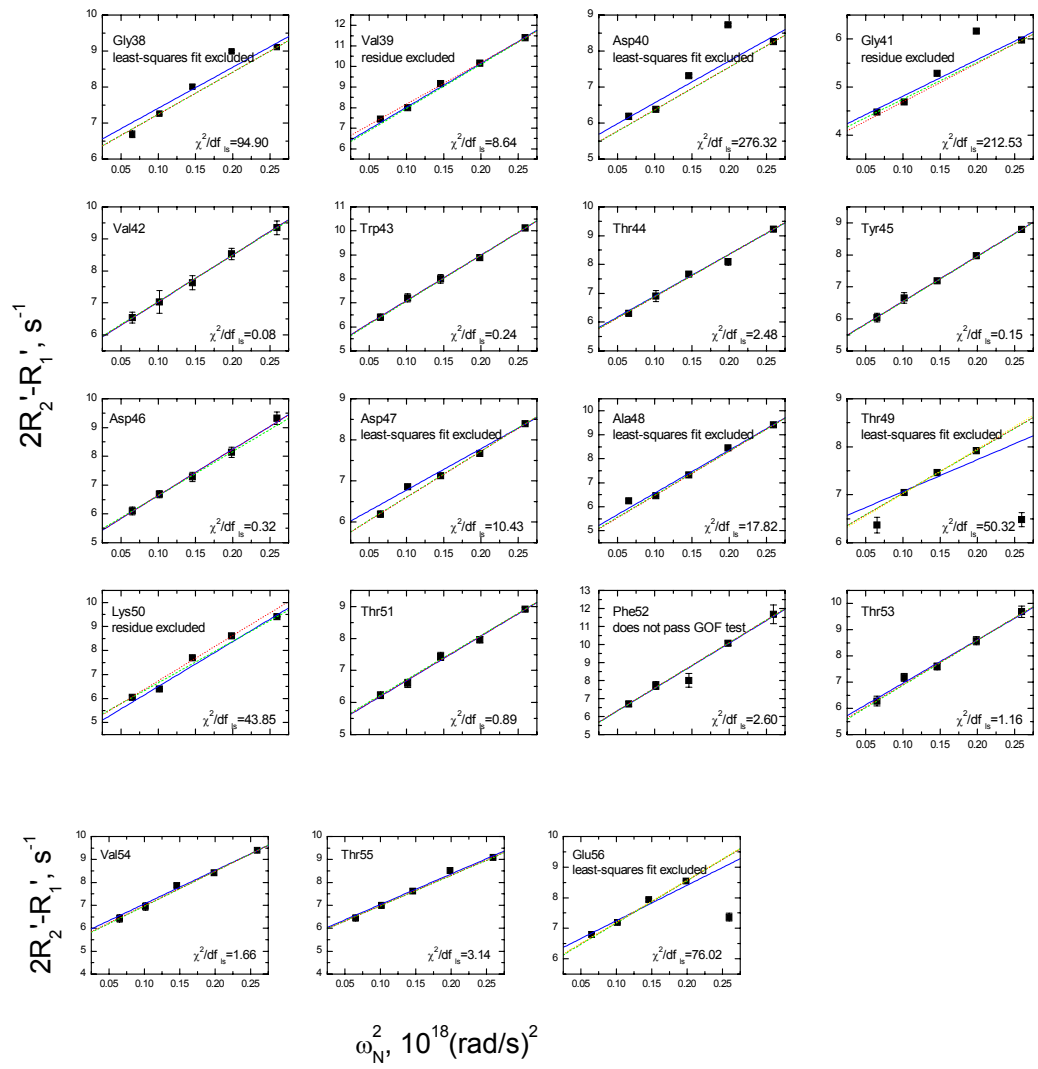


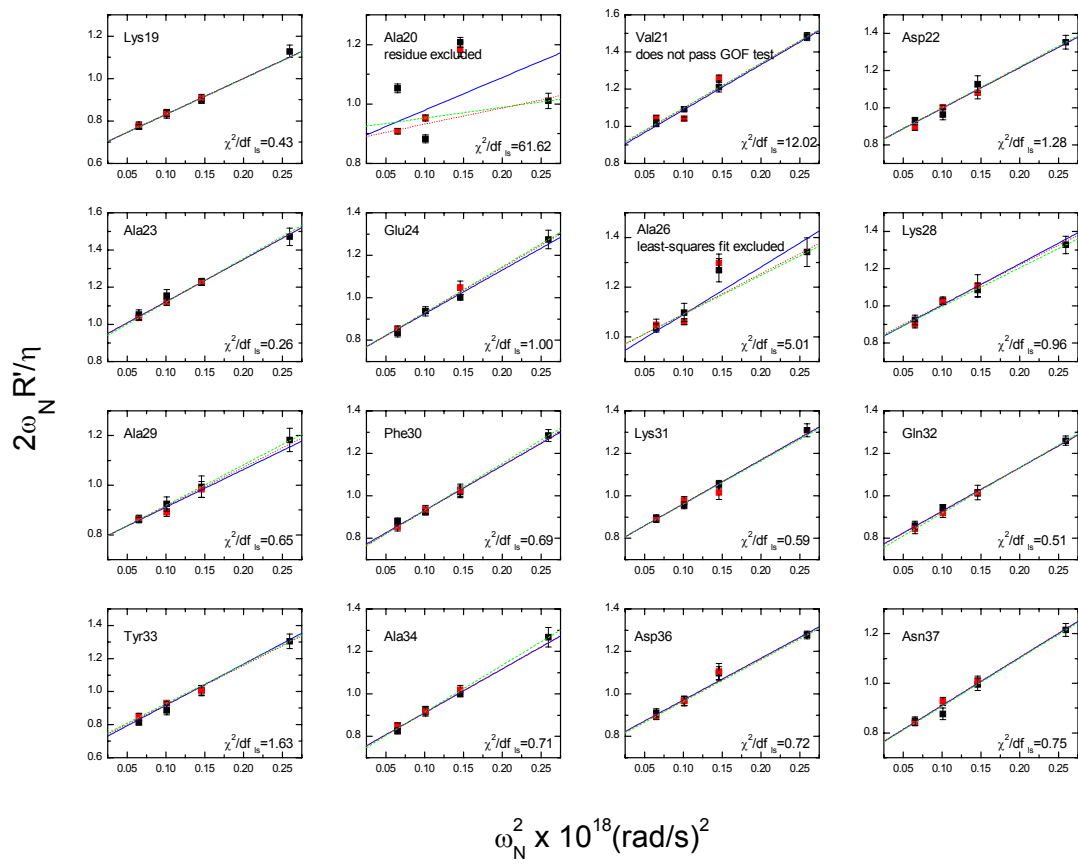
symmetric to a good approximation, thus from this protein it is impossible to discern the effects of neglecting rhombicity of the diffusion tensor. A protein with significant rhombicity of overall rotational diffusion would be needed for such an analysis.

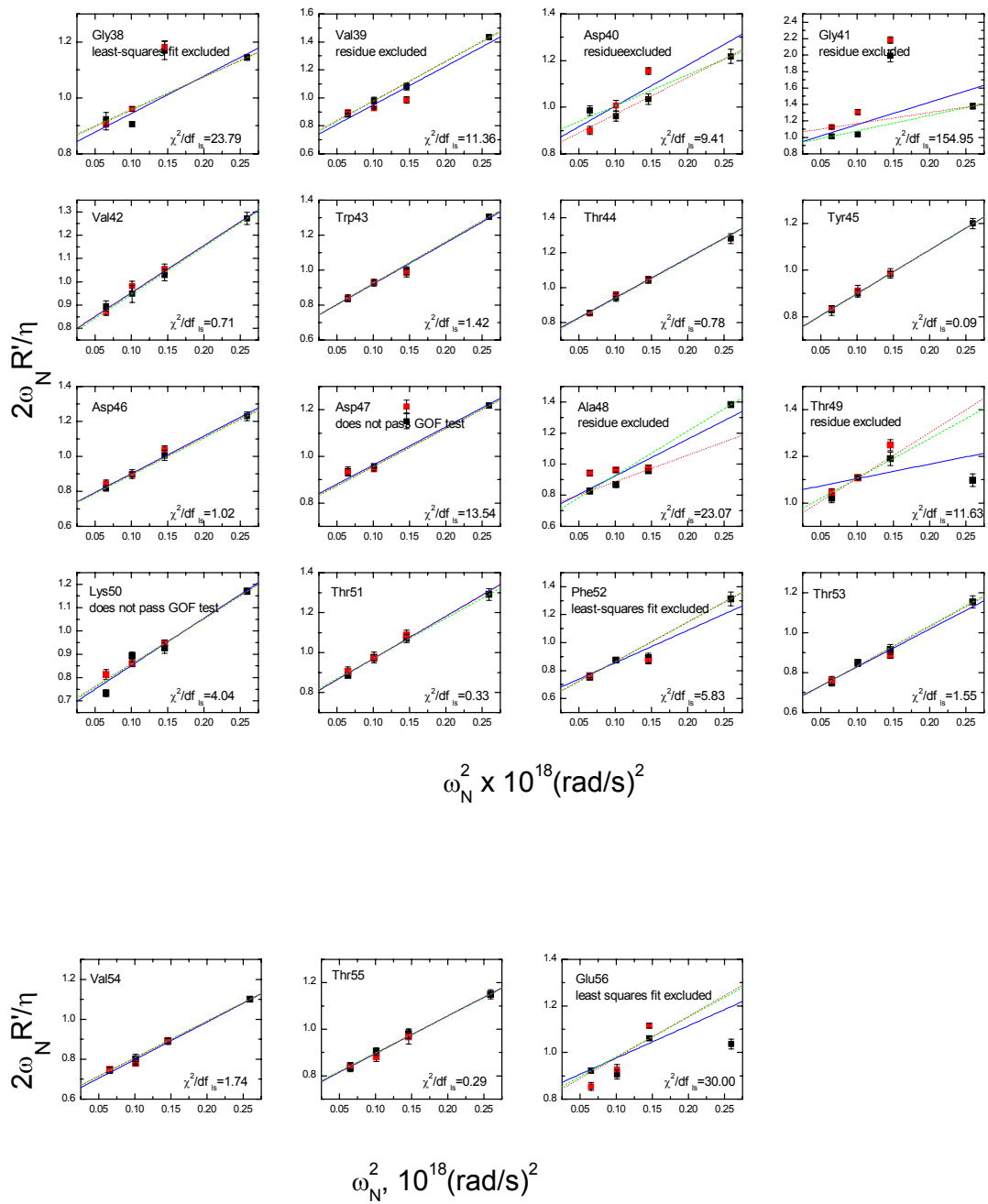
Here I have presented a thorough analysis of the variability in the  $^{15}\text{N}$  CSA in GB3. Still in question is the variability of the internuclear  $^{15}\text{N}$ - $^1\text{H}$  bond from residue to residue within the protein. Since the  $^{15}\text{N}$  CSA values in all the methods presented in Chapter 7 are determined via the dipolar term  $d$ , they therefore depend on our knowledge of the NH-bond length. Here, we have, as is customary, assumed a uniform value of the NH bond, so site-to-site variations in  $r_{\text{HN}}$  will necessarily affect the determined CSA values (a deviation in the bond length by  $\delta r_{\text{HN}}$  will introduce an error in the CSA value of the order of  $3(\delta r_{\text{HN}}/r_{\text{HN}})$ ). The direct-detected relaxation rates presented in Chapter 8 have the potential to help discern the effect of the variability in  $r_{\text{HN}}$ , since they are significantly less sensitive to the values of  $r_{\text{HN}}$ . Measurements at several field strengths will be necessary for a robust application of model-independent methods for CSA determination from these rates.

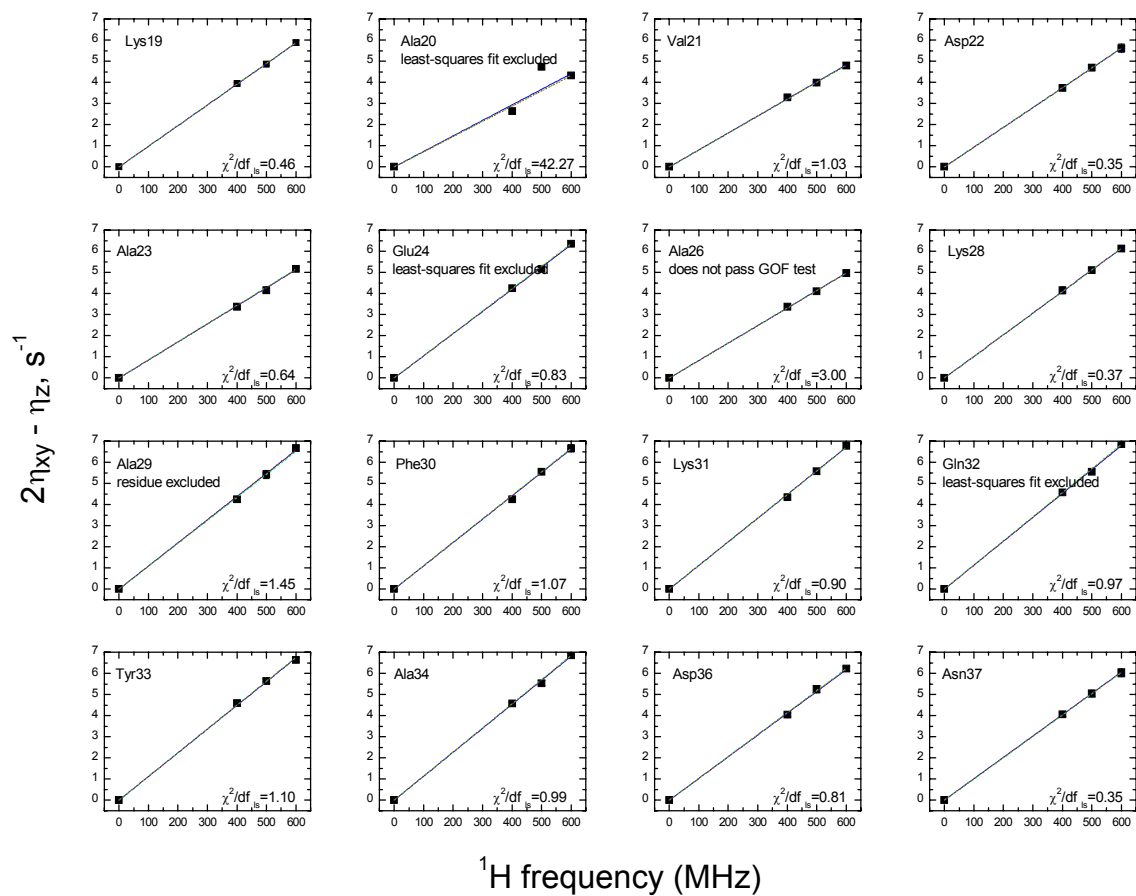
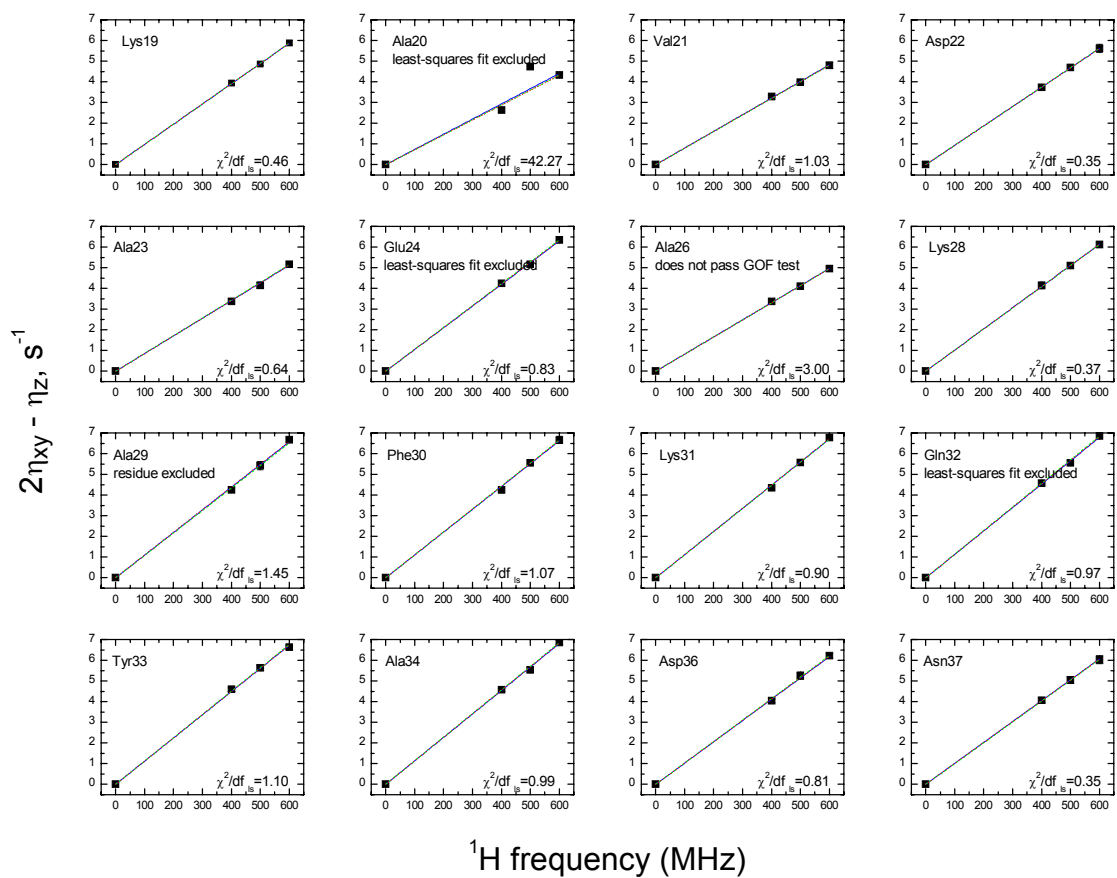
# Appendix A

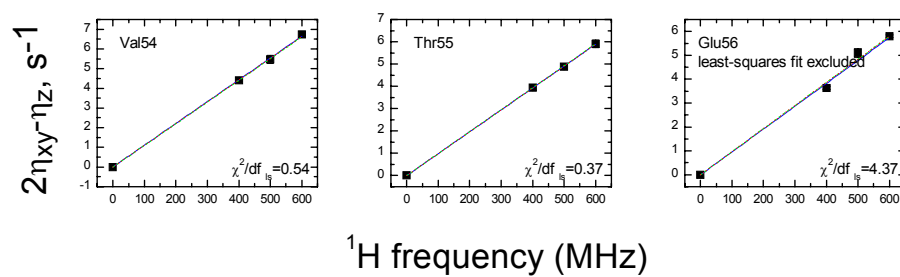
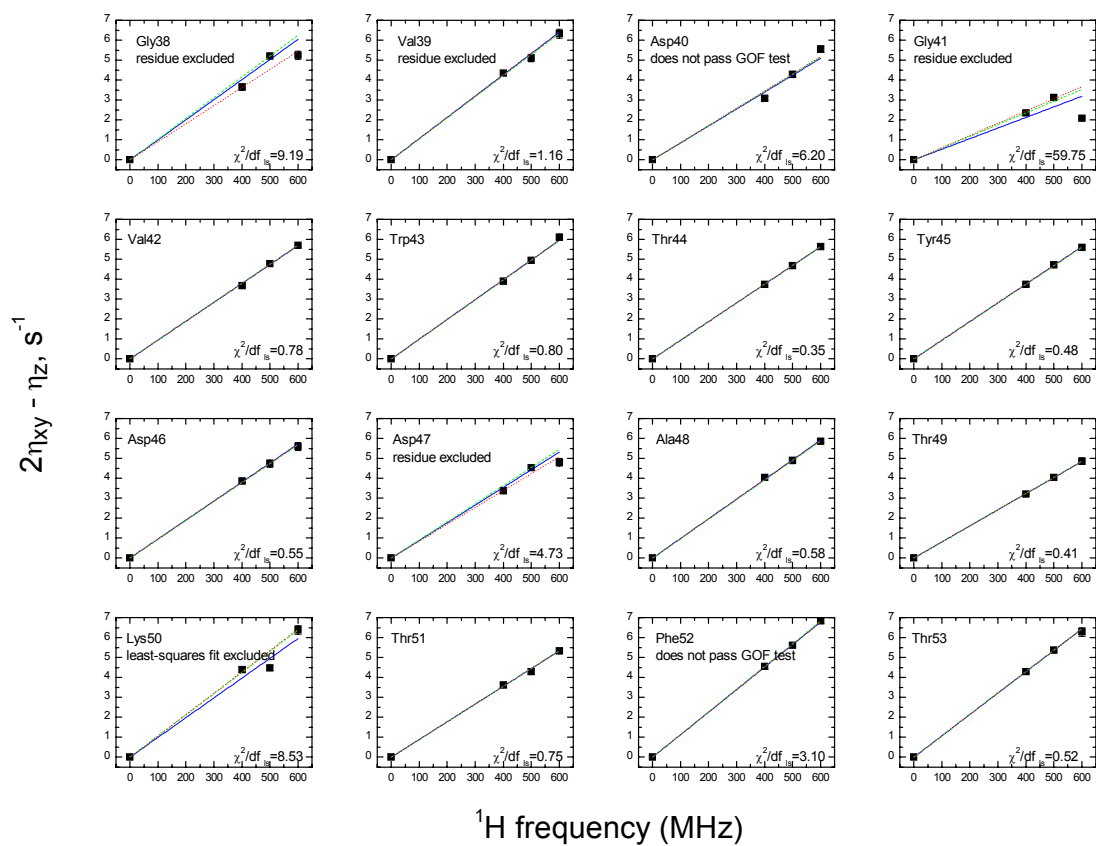


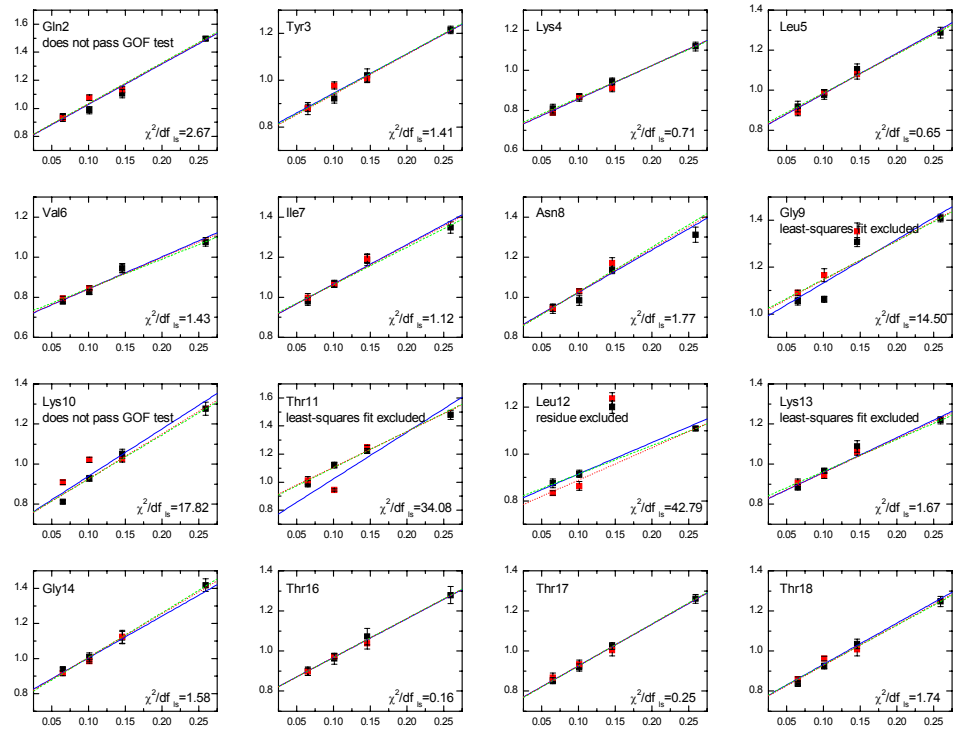
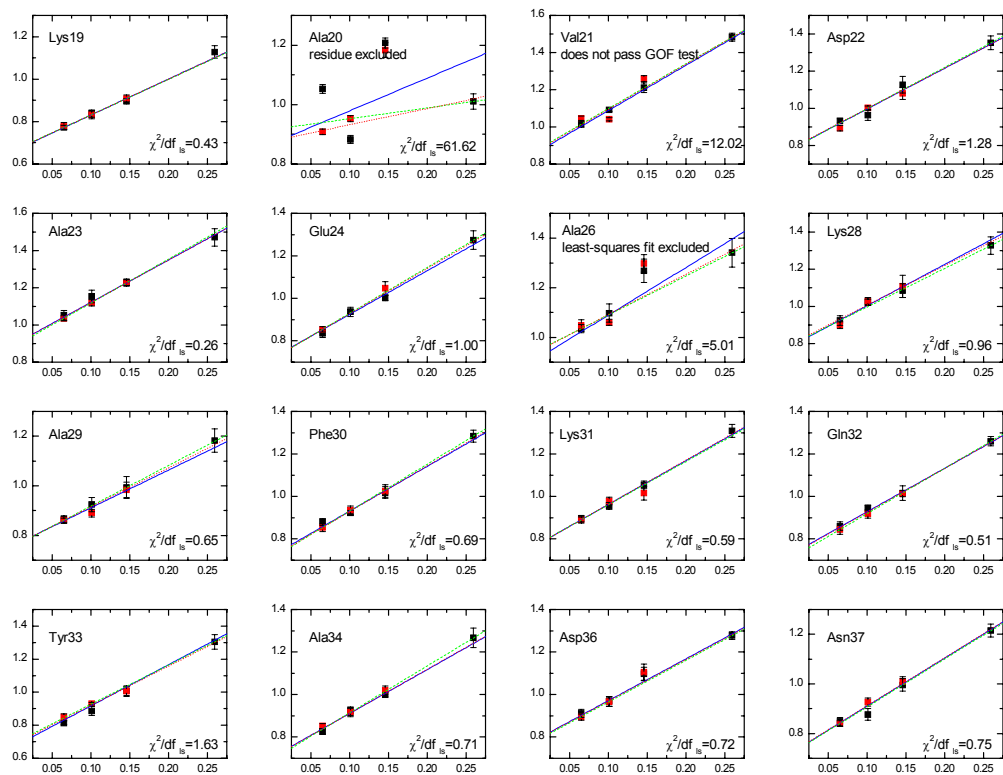




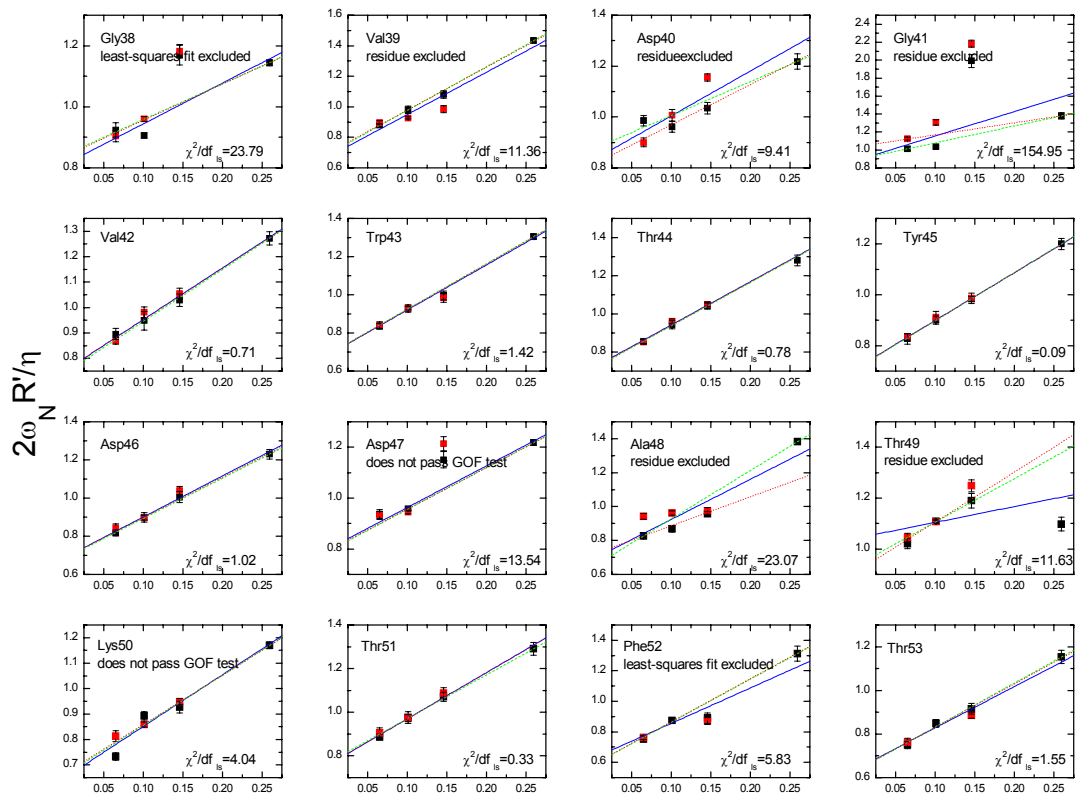




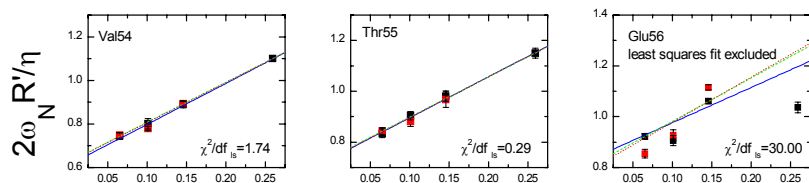


$2\omega_N R'/\eta$ 

 $\omega_N^2 \times 10^{18} (\text{rad/s})^2$ 
 $2\omega_N R'/\eta$ 

 $\omega_N^2 \times 10^{18} (\text{rad/s})^2$

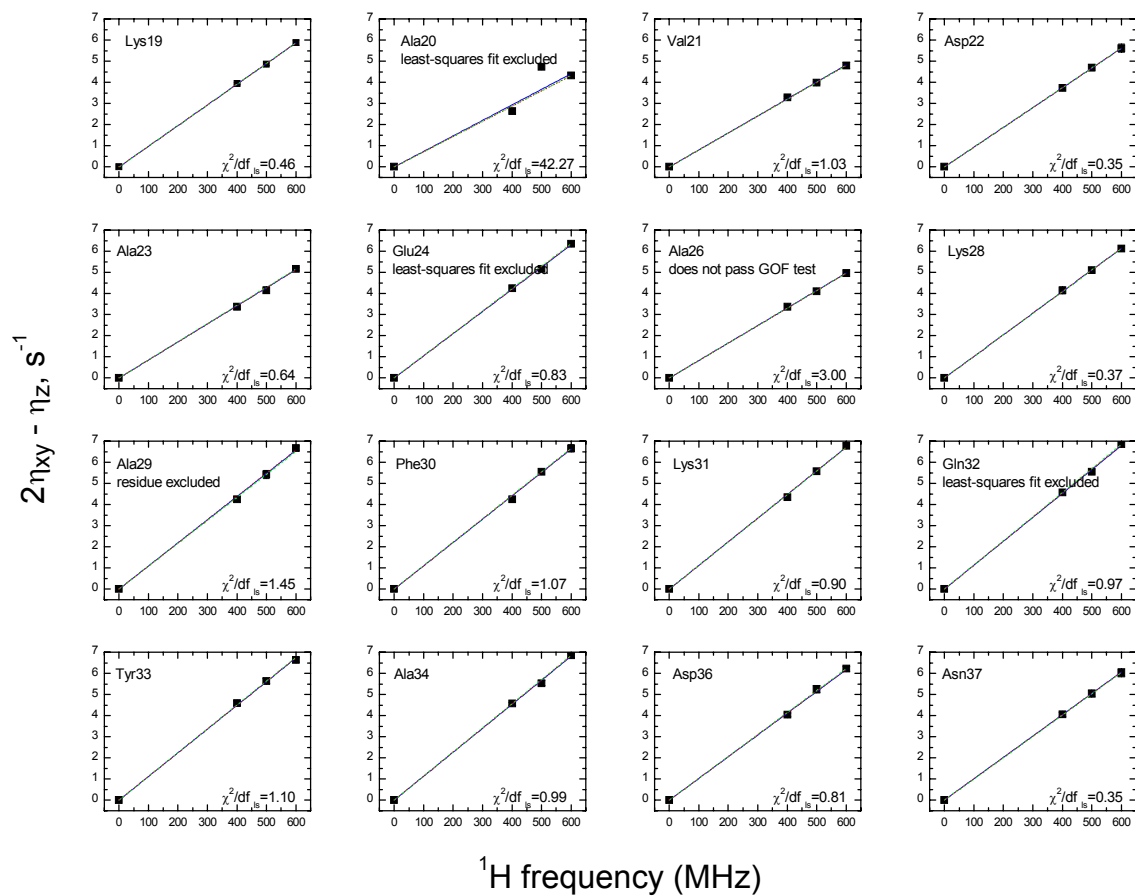
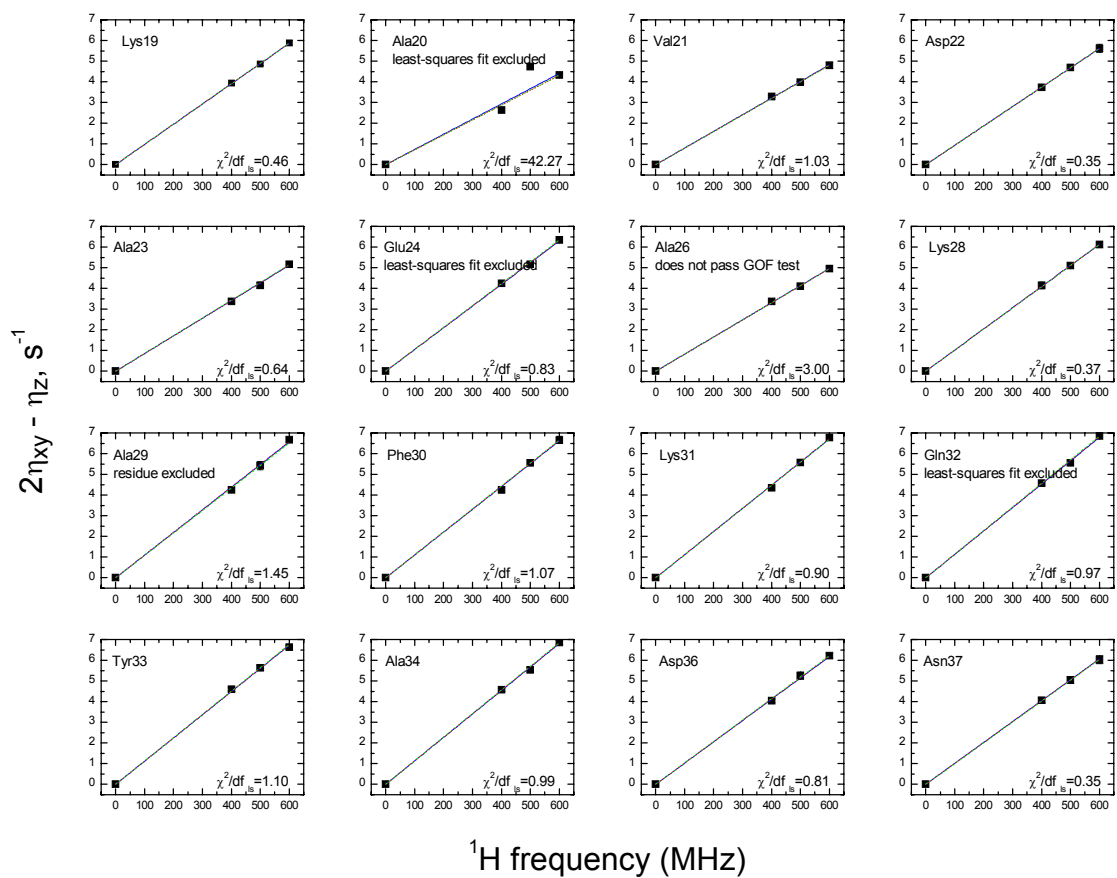


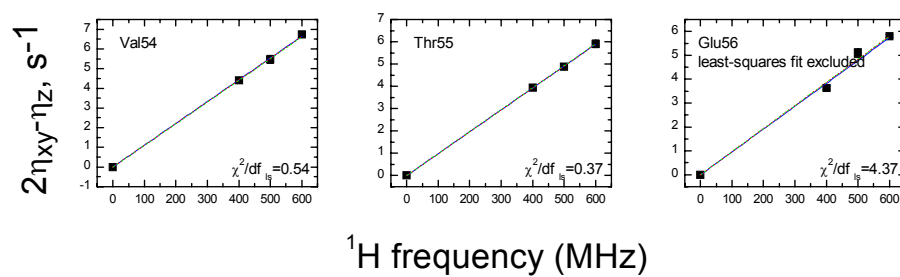
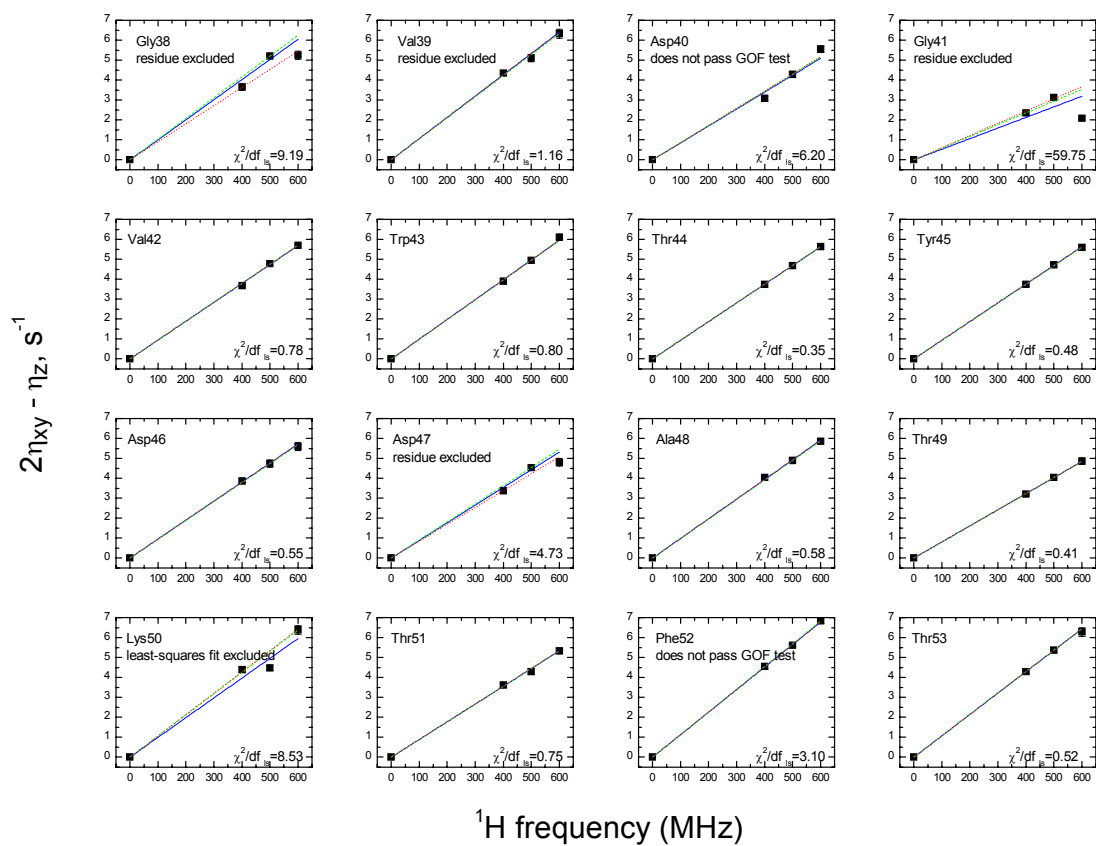


$$\omega_N^2 \times 10^{18} (\text{rad/s})^2$$



$$\omega_N^2, 10^{18} (\text{rad/s})^2$$







## Bibliography

1. Volkman, B.F., Lipson, D., Wemmer, D.E. & Kern, D. Two-state allosteric behavior in a single-domain signaling protein. *Science* **291**, 2429-33 (2001).
2. Lee, A.L., Kinnear, S.A. & Wand, A.J. Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. *Nat Struct Biol* **7**, 72-7 (2000).
3. Rabi, I.I. Space Quantization in a Gyating Magnetic Field. *Phys Rev* **51**, 652 (1937).
4. Pravia, M., Nelson, R.J., Weinstein, Y. & Cory, D.G. Observing Quantum Behavior in a Spin System by NMR. *Concepts Magn. Reson.* **11**, 225-238 (1999).
5. Ernst, R.R. Nuclear Magnetic Resonance Fourier Transform Spectroscopy, Nobel Lecture. 1-57 (Stockholm, Sweden, 1992).
6. Cory, D.G., Fahmy, A.F. & Havel, T.F. Ensemble quantum computing by NMR spectroscopy. *PNAS* **94**, 1634 (1997).
7. Gershenfeld, N. & Chuang, I. Bulk Spin-Resonance Quantum Computation. *Science* **275**, 350-356 (1997).
8. Bloembergen, N., Purcell, E.M. & Pound, R.V. Relaxation Effects in Nuclear Magnetic Resonance Absorption. *Phys Rev* **73**, 679 (1948).
9. Lindorff-Larsen, K., Best, R.B., Depristo, M.A., Dobson, C.M. & Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **433**, 128-32 (2005).
10. McCammon, J.A. & Harvey, S.C. *Dynamics of proteins and nucleic acids*, (Cambridge University Press, Cambridge, 1987).
11. Pfeiffer, S., Fushman, D. & Cowburn, D. Simulated and NMR derived backbone dynamics of a protein with significant flexibility: A comparison of spectral densities for the <math>\beta</math>-ARK PH domain. *J.Am.Chem.Soc.* **123**, 3021-36 (2001).
12. Nagy, A.M., Raicu, V. & Miller, R.J. Nonlinear optical studies of heme protein dynamics: Implications for proteins as hybrid states of matter. *Biochim Biophys Acta* **1749**, 148-72 (2005).
13. Austin, R.H., Beeson, K.W., Eisenstein, H., Frauenfelder, H. & Gunsalus, I.C. Dynamics of Ligand Binding to Myoglobin. *Biochemistry* **14**, 5355 (1975).
14. Clore, G.M. & Schwieters, C.D. Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: a unified picture of high probability, fast atomic motions in proteins. *J Mol Biol* **355**, 879-86 (2006).
15. Wang, T., Frederick, K.K., Igumenova, T.I., Wand, A.J. & Zuiderweg, E.R. Changes in calmodulin main-chain dynamics upon ligand binding revealed by cross-correlated NMR relaxation measurements. *J Am Chem Soc* **127**, 828-9 (2005).

16. Frauenfelder, H., Sligar, S.G. & Wolynes, P.G. The Energy Landscapes and Motions of Proteins. *Science* **254**, 1598-1603 (1991).
17. Kern, D., Eisenmesser, E.Z. & Wolf-Watz, M. Enzyme Dynamics During Catalysis Measured by NMR Spectroscopy. in *Methods in Enzymology* (Elsevier, 2005).
18. Eisenmesser, E.Z. et al. Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **438**, 117-121 (2005).
19. Tobi, D. & Bahar, I. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc Natl Acad Sci U S A* **102**, 18908-13 (2005).
20. Bahar, I. & Rader, A.J. Coarse-grained normal mode analysis in structural biology. *Curr Opin Struct Biol* **15**, 586-92 (2005).
21. Yang, L.W. & Bahar, I. Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. *Structure* **13**, 893-904 (2005).
22. Xu, C., Tobi, D. & Bahar, I. Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T $\leftrightarrow$ R2 transition. *J Mol Biol* **333**, 153-68 (2003).
23. Chennubhotla, C., Rader, A.J., Yang, L.W. & Bahar, I. Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. *Phys Biol* **2**, S173-80 (2005).
24. Bloch, F. Nuclear Induction. *Phys Rev* **70**, 460-474 (1946).
25. Xu, J., Millet, O., Kay, L.E. & Skrynnikov, N.R. A new spin probe of protein dynamics: nitrogen relaxation in  $^{15}\text{N}$ - $^2\text{H}$  amide groups. *J Am Chem Soc* **127**, 3220-9 (2005).
26. Yip, G.N. & Zuiderweg, E.R. A phase cycle scheme that significantly suppresses offset-dependent artifacts in the R2-CPMG  $^{15}\text{N}$  relaxation experiment. *J Magn Reson* **171**, 25-36 (2004).
27. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546-4559 (1982).
28. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *JACS* **104**, 4559-4570 (1982).
29. Clore, G.M. et al. Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc* **112**, 4989-4936 (1990).
30. Peng, J. & Wagner, G. Mapping spectral density functions using heteronuclear NMR relaxation measurements. *J Magn Reson* **94**, 82-100 (1992).
31. Farrow, N., Zhang, O., Forman-Kay, J. & Kay, L. Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry* **34**, 868-878 (1995).
32. Fushman, D., Ohlenschlager, O. & Rüterjans, H. Determination of the backbone mobility of ribonuclease T1 and its 2'GMP complex using molecular dynamics simulations and NMR relaxation data. *J Biomol Struct Dyn* **11**, 1377-402 (1994).

33. It is worth noting here that  $^{15}\text{N}$  spin rates are not sensitive to translational motion and the dipolar part of  $^{15}\text{N}$  autorelaxation rates ( $R_1$ ,  $R_2$ ) and cross-relaxation rates (steady state heteronuclear NOE) are not sensitive to motion that occurs along an axis parallel to the dipole-dipole interaction with the directly bound  $^1\text{H}$  nucleus. This renders these  $^{15}\text{N}$  relaxation rates insensitive to roll type motions of  $\alpha$ -helices.  $^{15}\text{N}$  CSA/dipolar cross-correlation rates (see Chapter 4) are in principal sensitive to these motions since they sample dynamics of the  $^{15}\text{N}$ - $^1\text{H}$  dipolar interaction projected onto the three orthogonal axes of the chemical shielding tensor.
34. Korzhnev, D.M., Billeter, M., Arseniev, A.S. & Orekhov, V., Y., NMR studies of Brownian tumbling and internal motions in proteins. *Progress NMR Spectroscopy* **38**, 197-266 (2001).
35. Wu, C.H., Ramamoorthy, A., Gierasch, L.M. & Opella, S.J. Simultaneous characterization of the amide  $^1\text{H}$  chemical shift,  $^1\text{H}$ - $^{15}\text{N}$  dipolar, and  $^{15}\text{N}$  chemical shift interaction tensors in a peptide bond by three-dimensional solid-state NMR spectroscopy. *J.Amer.Chem.Soc.* **117**, 6148-9 (1995).
36. Wishart, D.S., Sykes, B.D. & Richards, F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* **222**, 311-333 (1991).
37. Osapay, K., Theriault, Y., Wright, P.E. & Case, D.A. Solution structure of carbonmonoxy myoglobin determined from nuclear magnetic resonance distance and chemical shift constraints. *J Mol Biol* **244**, 183-97 (1994).
38. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* **13**, 289-302 (1999).
39. Skrynnikov, N.R., Dahlquist, F.W. & Kay, L.E. Reconstructing NMR spectra of "invisible" excited protein states using HSQC and HMQC experiments. *J Am Chem Soc* **124**, 12352-60 (2002).
40. Korzhnev, D.M. et al. Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586-90 (2004).
41. Harbison, G.S. et al. Chemical shift and  $^{15}\text{N}$ - $^{13}\text{C}$  dipolar tensors for the peptide bond in  $1\text{-}^{13}\text{C}$ -glycyl- $^{15}\text{N}$ -glycine hydrochloride monohydrate. *J Magn Reson* **60**, 79-82 (1984).
42.  $^{15}\text{N}$  relaxation rates are proportional to both the  $^{15}\text{N}$  CSA and the NH bond length, therefore assuming a particular uniform value of the bond length causes the  $^{15}\text{N}$  CSA to be uniformly scaled. To avoid confusion, all  $^{15}\text{N}$  CSA values reported herein have been scaled to their values assuming a NH bond length of 1.02 Å.
43. Tjandra, N., Szabo, A. & Bax, A. Protein backbone dynamics and  $^{15}\text{N}$  chemical shift anisotropy from quantitative measurement of relaxation interference effects. *J. Am. Chem. Soc.* **118**, 6986-6991 (1996).
44. Tjandra, N., Wingfield, P., Stahl, S. & Bax, A. Anisotropic rotational diffusion of perdeuterated HIV protease from  $^{15}\text{N}$  NMR relaxation measurements at two magnetic fields. *J Biomol NMR* **8**, 273-84 (1996).
45. Cornilescu, G. & Bax, A. Measurement of Proton, Nitrogen, and Carbonyl Chemical Shielding Anisotropies in a Protein Dissolved in a Dilute Liquid Crystalline Phase. *J. Am. Chem. Soc.* **122**, 10143-10154 (2000).

46. Boyd, J. & Redfield, C. Characterization of  $^{15}\text{N}$  Chemical Shift Anisotropy from Orientation-Dependent Changes to  $^{15}\text{N}$  Chemical Shifts in Dilute Bicelle Solutions. *J. Am. Chem. Soc.* **121**, 7441-2 (1999).
47. Walling, A.E., Pargas, R.E. & deDios, A.C. Chemical Shift Tensors in Peptides: A Quantum Mechanical Study. *J Phys Chem A* **101**, 7299-7303 (1997).
48. Poon, A., Birn, J. & Ramamoorthy, A. How Does an Amide- $^{15}\text{N}$  Chemical Shift Tensor Vary in Peptides. *Journal of Physical Chemistry B* **108**, 16577-16585 (2004).
49. Kuroki, S., Ando, S. & Ando, I. Hydrogen-bonding effect on  $^{15}\text{N}$  NMR chemical shifts of the glycine residue of oligopeptides in the solid state as studied by high-resolution solid-state NMR spectroscopy. *Journal of Molecular Structure* **240**, 19 (1990).
50. Fushman, D. & Cowburn, D. Model-independent analysis of  $^{15}\text{N}$  chemical shift anisotropy from NMR relaxation data. Ubiquitin as a test example. *J.Am.Chem.Soc.* **120**, 7109-10 (1998).
51. Fushman, D., Tjandra, N. & Cowburn, D. Direct measurement of  $^{15}\text{N}$  chemical shift anisotropy in solution. *J.Am.Chem.Soc.* **120**, 10947-52 (1998).
52. Fushman, D., Tjandra, N. & Cowburn, D. An approach to direct determination of protein dynamics from  $^{15}\text{N}$  NMR relaxation at multiple fields, independent of variable  $^{15}\text{N}$  chemical shift anisotropy and chemical exchange contributions. *J. Am. Chem. Soc.* **121**, 8577-8582 (1999).
53. Kroenke, C.D., Rance, M. & Palmer, A.G.I. Variability of the  $^{15}\text{N}$  chemical shift anisotropy in Escherichia coli ribonuclease H in solution. *J.Am.Chem.Soc.* **121**, 10119-25 (1999).
54. Damberg, P., Jarvet, J. & Graslund, A. Limited variations in  $^{15}\text{N}$  CSA magnitudes and orientations in ubiquitin are revealed by joint analysis of longitudinal and transverse NMR relaxation. *J Am Chem Soc* **127**, 1995-2005 (2005).
55. Damberg, P., Jarvet, J., Allard, P. & Graslund, A. Quantitative estimation of magnitude and orientation of the CSA tensor from field dependence of longitudinal relaxation rates. *J.Biomol.NMR* **15**, 27-37 (1999).
56. Loth, K., Pelupessy, P. & Bodenhausen, G. Chemical shift anisotropy tensors of carbonyl, nitrogen, and amide proton nuclei in proteins through cross-correlated relaxation in NMR spectroscopy. *J Am Chem Soc* **127**, 6062-8 (2005).
57. Kover, K.E. & Batta, G. Separating structure and dynamics in CSA/DD cross-correlated relaxation: a case study on trehalose and ubiquitin. *J Magn Reson* **150**, 137-46. (2001).
58. Canet, D., Barthe, P., Mutzenhardt, P. & Roumestand, C. A Comprehensive Analysis of Multifield  $^{15}\text{N}$  Relaxation Parameters in Proteins: Determination of  $^{15}\text{N}$  Chemical Shift Anisotropies. *J. Am. Chem. Soc.* **123**, 4567-4576 (2001).
59. Bouguet-Bonnet, S., Mutzenhardt, P., Roumestand, C. & Canet, D. The concept of effective correlation times for describing backbone motions in proteins. Part I. A residue-per-residue self-consistent analysis of multifield



- 15N relaxation parameters. *Concepts in Magnetic Resonance A* **24A**, 1-9 (2005).
60. Bouguet-Bonnet, S., Mutzenhardt, P., Roumestand, C. & Canet, D. The concept of effective correlation times for describing backbone motions in proteins. Part II. Tentative interpretation of the residue specific correlation time in terms of overall rotation-diffusion. *Concepts in Magnetic Resonance A* **24A**, 10-16 (2005).
  61. Ottiger, M., Delaglio, F. & Bax, A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson* **131**, 373-8 (1998).
  62. Meissner, A., Duus, J.o. & Sorensen, O.W. Spin-State-Selective Excitation. Application for E.COSY-Type Measurement of  $J_{HH}$  Coupling Constants. *J.Mag.Reson* **128**, 92-97 (1997).
  63. Luginbuhl, P., Pervushin, K.V., Iwai, H. & Wuthrich, K. Anisotropic molecular rotational diffusion in 15N spin relaxation studies of protein mobility. *Biochemistry* **36**, 7305-12 (1997).
  64. Fushman, D. & Cowburn, D. Studying protein dynamics with NMR relaxation. in *Structure, Motion, Interaction and Expression of Biological Macromolecules* (eds. Sarma, R. & Sarma, M.) 63-77 (Adenine Press, Albany, NY, 1998).
  65. Abragam, A. & Pound, R.V. Influence of Electric and Magnetic Fields on Angular Correlations. *Phys Rev* **92**, 943 (1953).
  66. Overhauser, A.W. Paramagnetic Relaxation in Metals. *Phys Rev* **89**, 689 (1953).
  67. Solomon, I. Relaxation Processes in a System of Two Spins. *Phys Rev* **99**, 559 (1955).
  68. Pines, D. & Slichter, C.P. Relaxation Times in Magnetic Resonance. *Phys Rev* **100**, 1014 (1950).
  69. Van Vleck, J.H. The Dipolar Broadening of Magnetic Resonance Lines in Crystals. *Phys Rev* **74**, 1168 (1948).
  70. Woessner, D. Nuclear spin relaxation in ellipsoids undergoing rotational brownian motion. *J. Chem. Phys* **37**, 647-654 (1962).
  71. Woessner, D.E., Snowden, B.S. & Meyer, G.H. Nuclear Spin-Lattice Relaxation in Axially Symmetric Ellipsoids with Internal Motion. *Journal of Chemical Physics* **50**, 719 (1969).
  72. Slichter, C.P. *Principles of Magnetic Resonance*, 1-655 (Springer-Verlag, Berlin, 1989).
  73. Fischer, M.W.F., Majumdar, A. & Zuiderweg, E.R.P. Protein NMR relaxation: theory, applications and outlook. *Prog. Nuc. Mag. Res. Spect.* **33**, 207-72 (1998).
  74. Canet, D. On the calculation of spectral density functions for spin interactions without axial symmetry. *Concepts Magn. Reson.* **10**, 291-297 (1998).
  75. Goldman, M. Interference effects in the relaxation of a pair of unlike spin-1/2 nuclei. *J.Magn.Reson.* **60**, 437-452 (1984).
  76. Wigner, E.P. *Group theory and its application to the quantum mechanics of atomic spectra*, 372 (Academic Press, New York, 1959).

77. Goldstein, H. *Classical Mechanics*, 146 (Addison-Wesley, 1980).
78. Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A. & Kay, L.E. Spectral density function mapping using  $^{15}\text{N}$  relaxation data exclusively. *J Biomol NMR* **6**, 153-62 (1995).
79. Ishima, R. & Nagayama, K. Protein backbone dynamics revealed by quasi spectral density function analysis of amide N- $^{15}\text{N}$  nuclei. *Biochemistry* **34**, 3162-3171 (1995).
80. Ernst, R.R. & Anderson, W.A. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Review of Scientific Instruments* **37**, 93-102 (1966).
81. Jeener, J., Meier, B.H., Bachmann, P. & Ernst, R.R. Investigation of exchange processes by two-dimensional NMR spectroscopy. *Journal of Chemical Physics* **71**, 4546-4553 (1979).
82. Aue, W.P., Bartholdi, E. & Ernst, R.R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *Journal of Chemical Physics* **64**, 2229-2246 (1976).
83. McIntosh, L.P. et al. Proton NMR measurements of bacteriophage T4 lysozyme aided by  $^{15}\text{N}$  isotopic labeling: structural and dynamic studies of larger proteins. *Proc Natl Acad Sci U S A* **84**, 1244-8 (1987).
84. Kumar, A., Ernst, R.R. & Wuthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Commun* **95**, 1-6 (1980).
85. Wuthrich, K., G. W., G. W. & W. B. Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance. *J.Mol.Biol.* **155**, 311-319 (1982).
86. Zheng, L., Fischer, M. & Zuiderweg, E. Study of protein dynamics in solution by measurement of  $^{13}\text{C}^{\alpha}$ - $^{13}\text{CO}$  NOE and  $^{13}\text{CO}$  longitudinal relaxation. *J.Biomol.NMR* **7**, 157-162 (1996).
87. Ulmer, T.S., Campbell, I.D. & Boyd, J. Amide proton relaxation measurements employing a highly deuterated protein. *J. Magn. Reson.* **166**, 190-201 (2004).
88. Buck, M. et al. Structural determinants of protein dynamics: analysis of  $^{15}\text{N}$  NMR relaxation measurements for main-chain and side-chain nuclei of hen egg white lysozyme. *Biochemistry* **34**, 4041-4055 (1995).
89. Mikhailov, D., Daragan, V.A. & Mayo, K.H. Lysine side-chain dynamics derived from  $^{13}\text{C}$ -multiplet NMR relaxation studies on di- and tripeptides. *J Biomol NMR* **5**, 397-410 (1995).
90. Stone, M.J., Chandrasekhar, K., Holmgren, A., Wright, P.E. & Dyson, H.J. Comparison of backbone and tryptophan side-chain dynamics of reduced and oxidized Escherichia coli thioredoxin using  $^{15}\text{N}$  NMR relaxation measurements. *Biochemistry* **32**, 426-35 (1993).
91. Millet, O., Mittermaier, A., Baker, D. & Kay, L.E. The effects of mutations on motions of side-chains in protein L studied by H-2 NMR dynamics and scalar couplings. *J Mol Biol* **329**, 551-563 (2003).

92. Ernst, R.H., Bodenhausen, G. & Wokaun, A. *Principles of nuclear magnetic resonance in one and two dimensions*, 610 (Clarendon Press, Oxford, 1987).
93. Carr, H.Y. & Purcell, E.M. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys.Rev.* **94**, 630-638 (1954).
94. Meiboom, S. & Gill, D. Modified spin-echo method for measuring nuclear relaxation times. *Rev.Sci.Instrum.* **29**, 688-691 (1958).
95. Grzesiek, S. & Bax, A. The importance of not saturating H<sub>2</sub>O in protein NMR -- application to sensitivity enhancement and NOE measurements. *J.Am.Chem.Soc.* **115**, 12593-4 (1993).
96. Jones, J.A. Optimal sampling strategies for the measurement of relaxation times in proteins. *J.Mag.Res.* **126**, 283-6 (1997).
97. Skelton, N. et al. Practical aspects of 2-dimensional proton-detected <sup>15</sup>N spin relaxation measurements. *J. Magn. Reson.* **B102**, 253-264 (1993).
98. Ghose, R., Huang, K. & Prestegard, J.H. Measurement of Cross Correlation between Dipolar Coupling and Chemical Shift Anisotropy in the Spin Relaxation of <sup>13</sup>C, <sup>15</sup>N-Labeled Proteins. *J.Mag.Res.* **135**, 487-499 (1998).
99. Ghose, R. & Prestegard, J.H. Improved estimation of CSA-dipolar coupling cross-correlation rates from laboratory-frame relaxation experiments. *J Magn Reson* **134**, 308-14 (1998).
100. Fushman, D. & Cowburn, D. Nuclear magnetic resonance relaxation in determination of residue-specific <sup>15</sup>N chemical shift tensors in proteins in solution: protein dynamics, structure, and applications of transverse relaxation optimized spectroscopy. in *Methods in Enzymology*, Vol. 339 (eds. James, T., Schmitz, U. & Doetsch, V.) 109-126 (2001).
101. Kroenke, C.D., Loria, J.P., Lee, L.K., Rance, M. & Palmer, A.G.I. Longitudinal and Transverse <sup>1</sup>H-<sup>15</sup>N Dipolar/<sup>15</sup>N Chemical Shift Anisotropy Relaxation Interference: Unambiguous Determination of Rotational Diffusion Tensors and Chemical Exchange Effects in Biological Macromolecules. *J. Am. Chem. Soc.* **120**, 7905 -7915 (1998).
102. Boisbouvier, J., Brutscher, B., Simorre, J.-P. & Marion, D. <sup>13</sup>C spin relaxation measurements in RNA: Sensitivity and resolution improvement using spin-state selective correlation experiments. *J. Biomol. NMR* **14**, 241-252 (1999).
103. Dayie, K.T., Brodsky, A.S. & Williamson, J.R. Base flexibility in HIV-2 TAR RNA mapped by solution (<sup>15</sup>N), (<sup>13</sup>C) NMR relaxation. *J Mol Biol* **317**, 263-78. (2002).
104. Tessari, M., Mulder, F.A.A., Boelens, R. & Vuister, G.W. Determination of amide proton CSA in <sup>15</sup>N labeled proteins using <sup>1</sup>H CSA/<sup>15</sup>N-<sup>1</sup>H dipolar and <sup>15</sup>N CSA/<sup>15</sup>N-<sup>1</sup>H dipolar cross-correlation rates. *J.Mag.Res.* **127**, 128-3 (1997).
105. Brutscher, B., Skrynnikov, N.R., Bremi, T., BrYschweiler, R. & Ernst, R.R. Quantitative investigation of dipole-CSA cross-correlated relaxation by ZQ/DQ spectroscopy. *J Magn Reson* **130**, 346-51 (1998).
106. Carlomagno, T., Maurer, M., Hennig, M. & Griesinger, C. Ubiquitin backbone motion studied via NH<sup>N</sup>-C'<sup>Ca</sup> dipolar-dipolar and C'-C'<sup>Ca</sup>/NH<sup>N</sup>

- CSA-dipolar cross-correlated relaxation. *J. Amer. Chem. Soc.* **122**, 5105-5113 (2000).
107. Batta, G., Kover, K.E. & Kowalewski, J. A comparison of 1D and 2D (unbiased) experimental methods for measuring CSA/DD cross-correlated relaxation. *J Magn Reson* **136**, 37-46. (1999).
  108. Pelupessy, P., Espallargas, G.M. & Bodenhausen, G. Symmetrical reconversion: measuring cross-correlation rates with enhanced accuracy. *J Magn Reson* **161**, 258-64 (2003).
  109. Carlomagno, T. & Griesinger, C. Errors in the measurement of cross-correlated relaxation rates and how to avoid them. *J Magn Reson* **144**, 280-7 (2000).
  110. Brutscher, B. Principles and applications of cross-correlated relaxation in biomolecules. *Concepts Magn. Reson.* **12**, 207-229 (2000).
  111. Kay, L.E., Nicholson, L.K., Delaglio, F., Bax, A. & Torchia, D. Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear  $T_1$  and  $T_2$  values in proteins. *J. Magn. Reson.* **97**, 359-375 (1992).
  112. Vasos, P.R., Hall, J.B. & Fushman, D. Spin-State Selection for Increased Confidence in Cross-Correlation Rates Measurements. *J. Biomol. NMR* **31**, 149-54 (2005).
  113. Sorensen, M.D., Meissner, A. & Sorensen, O.W. Spin-state-selective coherence transfer via intermediate states of two-spin coherence in IS spin systems: Application to E.COSY-type measurement of J coupling constants. *J. Biomol. NMR* **10**, 181-186 (1997).
  114. Andersson, P., Annala, A. & Otting, G. An alpha/beta-HSQC-alpha/beta experiment for spin-state selective editing of IS cross peaks. *J Magn Reson* **133**, 364-7 (1998).
  115. Bouguet-Bonnet, S., Mutzenhardt, P. & Canet, D. An alternative spin-state-selective pulse sequence element. *Magnetic Resonance in Chemistry* **41**, 1030-1033 (2003).
  116. Wang, L., Kurochkin, A.V. & Zuiderweg, E.R. An iterative fitting procedure for the determination of longitudinal NMR cross-correlation rates. *J Magn Reson* **144**, 175-85. (2000).
  117. Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. *Numerical Recipes in C*, (Cambridge University Press, NY, 1992).
  118. Barchi, J.J., Jr., Grasberger, B., Gronenborn, A.M. & Clore, G.M. Investigation of the backbone dynamics of the IgG-binding domain of streptococcal protein G by heteronuclear two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  nuclear magnetic resonance spectroscopy. *Protein Sci* **3**, 15-21 (1994).
  119. Schurr, J.M., Babcock, H.P. & Fujimoto, B.S. A test of the model-free formulas: Effects of anisotropic rotational diffusion and dimerization. *J. Magnetic Resonance* **B105**, 211-224 (1994).
  120. Walker, O., Varadan, R. & Fushman, D. Efficient and accurate determination of the overall rotational diffusion tensor of a molecule from  $^{15}\text{N}$  relaxation data using computer program ROTDIF. *J. Magn. Reson.* **168**, 336-345 (2004).

121. Ghose, R., Fushman, D. & Cowburn, D. Determination of the Rotational Diffusion Tensor of Macromolecules in Solution from NMR Relaxation Data with a Combination of Exact and Approximate Methods - Application to the Determination of Interdomain Orientation in Multidomain Proteins. *J.Magn.Reson.* **149**, 214-217 (2001).
122. Fushman, D. Surface fractality of proteins from theory and NMR data. *J.Biomol.Struct. & Dynam.* **7**, 1333-1344 (1990).
123. Tirado, M.M. & de la Torre, J.G. Rotational dynamics of rigid, symmetric top macromolecules. Application to circular cylinders. *J. Chem. Phys.* **73**, 1986-1993 (1980).
124. Copie, V. et al. Solution structure and dynamics of linked cell attachment modules of mouse fibronectin containing the RGD and synergy regions: comparison with the human fibronectin crystal structure. *J.Mol.Biol.* **277**, 663-682 (1998).
125. Despa, F., Fernandez, A. & Berry, R.S. Dielectric Modulation of Biological Water. *Phys Rev Lett* **93**, 228104 (2004).
126. Byron, O. Hydrodynamic bead modeling of biological macromolecules. *Numerical Computer Methods, Part C* **321**, 278-304 (2000).
127. de la Torre, J.G. & Bloomfield, V.A. Hydrodynamic properties of complex, rigid, biological macromolecules. Theory and applications. *Q. Rev. Biophys.* **14**, 81-139 (1981).
128. de la Torre, J.G., Navarro, S., Martinez, M.C.L., Diaz, F.G. & Cascales, J.J.L. HYDRO: A computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys. J.* **67**, 530-531 (1994).
129. Burling, F.T., Weis, W.I., Flaherty, K.M. & Brunger, A. Direct Observation of Protein Solvation and Discrete Disorder with Experimental Crystallographic Phases. *Science* **271**, 72-77 (1996).
130. Garcia de la Torre, J., Huertas, M.L. & Carrasco, B. HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **B147**, 138-46 (2000).
131. Garcia de la Torre, J., Bernado, P. & Pons, M. Hydrodynamic models and computational methods for NMR relaxation. *Methods Enzymol* **394**, 419-30 (2005).
132. Eisenmesser, E.Z., Bosco, D.A., Akke, M. & Kern, D. Enzyme dynamics during catalysis. *Science* **295**, 1520-3. (2002).
133. Kern, D. & Zuiderweg, E.R. The role of dynamics in allosteric regulation. *Current Opinion in Structural Biology* **13**, 748-757 (2003).
134. Rousseau, F. & Schymkowitz, J. A systems biology perspective on protein structural dynamics and signal transduction. *Current Opinion in Structural Biology* **15**, 23-30 (2005).
135. Fersht, A. *Structure and Mechanism in Protein Science*, (W.H. Freeman, New York, 2000).
136. Akke, M. & Palmer, A.G., III. Monitoring macromolecular motions on microsecond-millisecond time scales by R(1rho) - R(1) constant-relaxation-time NMR spectroscopy. *J. Am. Chem. Soc.* **118**, 911-912 (1996).

137. Loria, J.P., Rance, M. & Palmer, A.G.I. A Relaxation-Compensated Carr-Purcell-Meiboom-Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *J. Am. Chem. Soc.* **121**, 2331-2 (1999).
138. Wang, C., Grey, M.J. & Palmer, A.G.I. CPMG sequences with enhanced sensitivity to chemical exchange. *J. Biomol. NMR* **21**, 361-6 (2001).
139. Butterwick, J.A. et al. Multiple time scale backbone dynamics of homologous thermophilic and mesophilic ribonuclease HI enzymes. *J Mol Biol* **339**, 855-71 (2004).
140. Wang, C., Karpowich, N., Hunt, J.F., Rance, M. & Palmer, A.G. Dynamics of ATP-binding cassette contribute to allosteric control, nucleotide binding and energy transduction in ABC transporters. *J Mol Biol* **342**, 525-37 (2004).
141. Blackledge, M.J. et al. Conformational backbone dynamics of the cyclic decapeptide antamanide. Application of a new multiconformational search algorithm based on NMR data. *Biochemistry* **32**, 10960-74 (1993).
142. Mandel, A.M., Akke, M. & Palmer, A.G., 3rd. Dynamics of ribonuclease H: temperature dependence of motions on multiple time scales. *Biochemistry* **35**, 16009-23 (1996).
143. Fushman, D., Cahill, S. & Cowburn, D. The main chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: Analysis of <sup>15</sup>N relaxation with monomer/dimer equilibration. *J. Mol. Biol.* **266**, 173-194 (1997).
144. Camarero, J.A. et al. Rescuing a destabilized protein fold through backbone cyclization. *J. Mol. Biol.* **308**, 1045-62 (2001).
145. Lian, L.Y., Derrick, J.P., Sutcliffe, M.J., Yang, J.C. & Roberts, G.C. Determination of the solution structures of domains II and III of protein G from *Streptococcus* by <sup>1</sup>H nuclear magnetic resonance. *J Mol Biol* **228**, 1219-34. (1992).
146. Millet, O., Loria, J., Kroenke, C.D., Pons, M. & Palmer, A.G. The static magnetic field dependence of chemical exchange linebroadening defines the NMR chemical shift time scale. *J.Am.Chem.Soc.* **122**, 2867-2877 (2000).
147. Bernado, P. & Blackledge, M. Anisotropic small amplitude Peptide plane dynamics in proteins from residual dipolar couplings. *J Am Chem Soc* **126**, 4907-20 (2004).
148. Bouvignies, G. et al. Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc Natl Acad Sci U S A* **102**, 13885-90 (2005).
149. Clore, G.M. & Schwieters, C.D. Amplitudes of protein backbone dynamics and correlated motions in a small alpha/beta protein: correspondence of dipolar coupling and heteronuclear relaxation measurements. *Biochemistry* **43**, 10678-91 (2004).
150. Kurita, J., Shimahara, H., Utsunomiya-Tate, N. & Tate, S. Measurement of <sup>15</sup>N chemical shift anisotropy in a protein dissolved in a dilute liquid crystalline medium with the application of magic angle sample spinning. *J Magn Reson* **163**, 163-73 (2003).

151. Lee, A.L. & Wand, A.J. Assessing bias in the determination of rotational correlation times of proteins by NMR relaxation. *J.Biomol.NMR* **13**, 101-112 (1999).
152. Hall, J.B. & Fushman, D. Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: Differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G. *J Biomol NMR* **27**, 261-275 (2003).
153. Rousseeuw, R. & Leroy, A.M. *Robust regression and outlier detection*, (John Wiley & Sons, New York, NY USA, 2003).
154. Draper, N.R. & Smith, H. *Applied regression analysis*, 709 (John Wiley & Sons, New York, 1998).
155. Mandel, A.M., Akke, M. & Palmer, A.G.I. Backbone dynamics of E. coli Ribonuclease HI: correlations with structure and function in an active enzyme. *J. Molecular Biology* **246**, 144-163 (1995).
156. Tjandra, N., Feller, S.E., Pastor, R.W. & Bax, A. Rotational diffusion anisotropy of human ubiquitin from <sup>15</sup>N NMR relaxation. *J. Am. Chem. Soc.* **117**, 12562-12566 (1995).
157. Case, D.A., Dyson, H.J. & Wright, P.E. Use of chemical shifts and coupling constants in nuclear magnetic resonance structural studies on peptides and proteins. *Methods Enzymol* **239**, 392-416 (1994).
158. Case, D. Use of chemical shifts and their anisotropies in biomolecular structure determination. *Current Opinions Structural Biology* **XXX**, in press (1998).
159. Cornilescu, G., Marquardt, J.L., Ottiger, M. & Bax, A. Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. *J. Am. Chem. Soc.* **120**, 6836-6837 (1998).
160. Korzhnev, D.M., Orekhov, V.Y. & Arseniev, A.S. Model-free approach beyond the borders of its applicability. *J Magn Reson* **127**, 184-91 (1997).
161. Fushman, D. & Cowburn, D. The effect of noncollinearity of 15N-1H dipolar and 15N CSA tensors and rotational anisotropy on 15N relaxation rates, CSA/DD cross correlation, and TROSY. *J. Biomol. NMR* **13**, 139-147 (1999).
162. Scheurer, C. et al. Effect of Dynamics and Environment on 15N Chemical Shielding Anisotropy in Proteins. A Combination of Density Functional Theory, Molecular Dynamics Simulation, and NMR Relaxation. *Journal of the American Chemical Society* **121**, 4242-4251 (1999).
163. Bremi, T. & Bruschweiler, R. Locally anisotropic internal polypeptide backbone dynamics by NMR relaxation. *J.Amer.Chem.Soc.* **119**, 6672-73 (1997).
164. Lienin, S.F., Bremi, T., Brutscher, B., Bruschweiler, R. & Ernst, R.R. Anisotropic intramolecular backbone dynamics of ubiquitin characterized by NMR relaxation and MD computer simulation. *J.Amer.Chem.Soc.* **120**, 9870-9 (1998).
165. Case, D.A. Calculations of NMR dipolar coupling strengths in model peptides. *J. Biomol. NMR* **15**, 95 (1999).

166. Bertini, I., Felli, I.C., Kummerle, R., Moskau, D. & Pierattelli, R.  $^{13}\text{C}$ - $^{13}\text{C}$  NOESY: an attractive alternative for studying large macromolecules. *J Am Chem Soc* **126**, 464-5 (2004).
167. Oh, B.H., Westler, W.M., Darba, P. & Markley, J.L. Protein carbon- $^{13}\text{C}$  spin systems by a single two-dimensional nuclear magnetic resonance experiment. *Science* **240**, 908-11 (1988).
168. Bermel, W., Bertini, I., Felli, I.C., Kummerle, R. & Pierattelli, R.  $^{13}\text{C}$  direct detection experiments on the paramagnetic oxidized monomeric copper, zinc superoxide dismutase. *J Am Chem Soc* **125**, 16423-9 (2003).
169. Bertini, I., Fragai, M., Lee, Y.M., Luchinat, C. & Terni, B. Paramagnetic metal ions in ligand screening: the Co(II) matrix metalloproteinase 12. *Angew Chem Int Ed Engl* **43**, 2254-6 (2004).
170. Eletsky, A., Moreira, O., Kovacs, H. & Pervushin, K. A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using  $^{13}\text{C}$  spectroscopy. *J Biomol NMR* **26**, 167-79 (2003).
171. Vogeli, B., Kovacs, H. & Pervushin, K. Measurements of side-chain  $^{13}\text{C}$ - $^{13}\text{C}$  residual dipolar couplings in uniformly deuterated proteins. *J Am Chem Soc* **126**, 2414-20 (2004).
172. Boyd, J., Mal, T.K., Soffe, N. & Campbell, I.D. The influence of a scalar-coupled deuterium upon the relaxation of a  $^{15}\text{N}$  Nucleus and its possible exploitation as a probe for side-chain interactions in proteins. *J Magn Reson* **124**, 61-71 (1997).
173. Fraenkel, G. & Asahi, Y. Isotope Shifts and Spin Coupling in the Proton Magnetic Resonance of Deuterated Ammonium Ions. *Journal of Chemical Physics* **44**, 4647-4649 (1966).
174. Munch, M., Hansen, A.E., Poul, E.H. & Bouman, T.D. Ab Initio Calculations of Deuterium Isotope Effects on Hydrogen and Nitrogen Nuclear Magnetic Shielding in the Hydrated Ammonium Ion. *Acta Chemica Scandinavica* **46**, 1065-1071 (1992).
175. Mennucci, B. & Martinez, J.M. How to Model Solvation of Peptides? Insights from a Quantum Mechanical and Molecular Dynamics Study of N-Methylacetamide. 2.  $^{15}\text{N}$  and  $^{17}\text{O}$  Nuclear Shielding in Water and Acetone. *J Phys Chem B* **109**, 9830-9838 (2005).
176. Wishart, D.S. et al.  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  chemical shift references in biomolecular NMR. *J. Biomol. NMR* **6**, 135-140 (1995).
177. Cho, C.H., Urquidi, J., Singh, S. & Robinson, G.W. Thermal offset viscosities of liquid  $\text{H}_2\text{O}$ ,  $\text{D}_2\text{O}$ , and  $\text{T}_2\text{O}$ . *J.Phys.Chem. B* **103**, 1991-1994 (1999).